Europäisches Patentamt
European Patent Office
Office européen des brevets

10/018009

EP00/6227

REC'D 15 SEP 2000
WIPO PCT

4

Bescheinigung Certificate Attestation

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

99201893.7

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

I.L.C. HATTEN-HECKMAN

CERTIFIED COPY OF PRIORITY DOCUMENT

DEN HAAG, DEN
THE HAGUE,
LA HAYE, LE 08/09/00

Europäisches Patentamt
European Patent Office
Office européen des brevets

Blatt 2 der Bescheinigung
Sheet 2 of the certificate
Page 2 de l'attestation

Anmeldung Nr.:
Application no.:
Demande n°: 99201893.7

Anmeldetag:
Date of filing:
Date de dépôt: 14/06/99

Anmelder:
Applicant(s):
Demandeur(s):
DSM N.V.
6411 TE Heerlen
NETHERLANDS

Bezeichnung der Erfindung:
Title of the invention:
Titre de l'invention:
Genes encoding enzymes in the biosynthesis of pimaricin and the application thereof

In Anspruch genommene Priorität(en) / Priority(ies) claimed / Priorité(s) revendiquée(s)

Staat: State: Pays:	Tag: Date: Date:	Aktenzeichen: File no. Numéro de dépôt:

Internationale Patentklassifikation:
International Patent classification:
Classification internationale des brevets:
C12N15/53, C12N15/52, C12N9/02, C12N9/04, C12P19/62

Am Anmeldetag benannte Vertragstaaten:
Contracting states designated at date of filing: AT/BE/CH/CY/DE/DK/ES/FI/FR/GB/GR/IE/IT/LI/LU/MC/NL/PT/SE
Etats contractants désignés lors du depôt:

Bemerkungen:
Remarks:
Remarques:

EP-2959 DSM N.V.

GENES ENCODING ENZYMES IN THE BIOSYNTHESIS OF PIMARICIN AND THE APPLICATION THEREOF

FIELD OF THE INVENTION

The invention relates to genes encoding enzymes which are fundamental in the biosynthesis of pimaricin. The invention further relates the application of said gene for modifying the biosynthesis of pimaricin, as well as for the biosynthesis of new compounds.

BACKGROUND OF THE INVENTION

Polyketides, such as pimaricin (in the literature also referred to as natamycin, see for its structure Fig. 3A), form a large and highly diverse group of natural products. Members of said group include compounds having antibacterial, antifungal, anticancer, antiparasitic and immunosuppressant activities. Despite their structural diversity, these metabolites are believed to be synthesized by micro-organisms by a common pathway in which units derived from acetate, propionate or butyrate are condensed onto the growing chain by a polyketide synthase (PKS), in a process resembling fatty acid biosynthesis, except that the β-keto function introduced at each elongation step may undergo all, part or none of a reductive cycle comprising β-ketoreduction, dehydration and enoylreduction. The structural variety in the group of polyketides arises from the choice of monomers, the extent of β-ketoreduction and dehydration, and the stereochemistry of each chiral center. Yet further diversity is produced by functionalization of the polyketide chain by the action of glycosylases, methyltransferases and oxidative enzymes.

Modification of complex biomolecules, such as polyketides, is an increasingly important way of obtaining

biologically active compounds with improved or altered properties. Currently, these modifications are usually introduced by chemical methods in a directed or random (e.g. in combinatorial chemistry) manner. A drawback of these chemical methods is that they are often performed under relatively harsh conditions. Furthermore, they lack selectivity and/or sensitivity. Particularly, in the case of complex biomolecules having multiple functionalized, reactive groups, precautions have to be taken in order to avoid undesired side reactions. These precautions include for instance the introduction of protective groups before a desired chemical conversion, involving two additional process steps, as the protective groups must be removed afterwards.

Bioconversion of simple organic compounds, i.e. compounds with no or single reactive centers, has been known for some time and finds application widely. Examples are the oxidation of long chain alkanes using alkane hydroxylation systems of *Pseudomonas*, and epoxidation of alkenes using enzyme systems from various micro-organisms. For the specific modifications required in the biosynthesis of complex molecules, like β-lactam antibiotics, polyketide antibiotics, anticancer agents, or peptide antibiotics, the large amounts of reactive groups present in these molecules are problematic for even the simplest treatments, like hydrolysis of specific bonds. More interesting treatments frequently completely destroy the molecule.

SUMMARY OF THE INVENTION

It is an object of the present invention to provide the means to perform specific conversions in complex biomolecules, in particular in polyketides, without applying the harsh conditions often related to chemical modifications. It is further an object of the invention that said conversions can be carried as part of a biosynthesis of said biomolecules, for instance in micro-organisms.

The invention provides the objective means in the form of three specific genes encoding for enzymes which facilitate specific oxidative conversions in the biosynthesis of pimaricin. Thus, the invention concerns a nucleotide sequence comprising SEQ ID NO. 5 (ORF1), a nucleotide sequence comprising SEQ ID NO. 7 (ORF2), a nucleotide sequence comprising SEQ ID NO. 9 (ORF3), and homologues or fragments thereof.

In the context of the present invention, a homologue gene sequence is defined as a sequence being for at least 75%, preferably at least 85%, more preferably at least 90% identical, while still encoding proteins having the desired enzymatic activity. Fragments of genes are, in accordance with the invention, defined as parts of the sequence having the indicated SEQ ID NO. having at least 90 nucleotides and encoding proteins having the desired enzymatic activity.

A homologue protein sequence is defined herein as a sequence being for at least 80%, preferably at least 90%, more preferably at least 95% identical, while still having the desired enzymatic activity. Fragments of proteins are, in accordance with the invention, defined as parts of the protein sequence having the indicated SEQ ID NO. having at least 30 amino acids and having the desired enzymatic activity.

Surprisingly, it has been found that the present genes can be expressed in different micro-organisms, which may be used in the biosynthesis of different biomolecules. It has further been found that said genes may be switched off (or knocked out) in the *Streptomyces* which is usually used for the biosynthesis of pimaricin. In this embodiment, not pimaricin is produced by said micro-organism, but a modified biomolecule. In addition, it has been found that said genes may be overexpressed in *Streptomyces*, leading to an increased production of said micro-organism in the biosynthesis of pimaricin.

DESCRIPTION OF THE DRAWINGS

Figure 1: Physical map of part of the Pimaricin biosynthetic cluster.

Genes: locations of the genes encoding polyketide synthases and oxidative genes involved in Pimaricin production (not drawn to scale);

Probes: 0.7 indicates the location of the 0.7 kb fragment used to identify the extent of polyketide synthase encoding regions; 3.3 indicates the location of the 3.3 kb fragment used in polyketide synthase gene disruption;

Cosmids: sizes and numbers of available cosmids covering the chromosomal region encompassing the oxidative genes.

Figure 2: Detailed physical map of the chromosomal regions including the oxidative genes.

Figure 3A: Molecular structure of Pimaricin.

Figure 3B: Molecular structures of Pimaricin derivatives with a reduced oxidation state of C4 and C5 and/or the carboxyl group at C12.

Figure 4: Molecular structures of Amphotericin B and Nystatin

DETAILED DESCRIPTION OF THE INVENTION

The functionality of the Pimaricin PKS associated genes was initially pursued by comparing their derived amino acid sequences with those present in public databases like EMBL, Genbank, NBRF/PIR, or Swissprot.

Surprisingly, ORF1 appeared to resemble cholesterol oxidases from several Streptomyces species. The close association of ORF1 with the Pimaricin PKS suggests an

oxidative step in Pimaricin tailoring. A cholesterol oxidase encoding gene has not been observed previously in a polyketide biosynthesis gene cluster.

Based in similar analyses, ORF2 and ORF3 resemble cytochrome P450 dependent monooxygenases from various sources. With respect to the biosynthesis of bioactive compounds, P450 dependent monooxygenases have been identified before in association with polyketide gene clusters, e.g. in the Erythromycin and Rapamycin biosynthesis gene clusters. Only in the Erythromycin case has the specific enzymatic action on Erythromycin precursor compounds been proven. Essentially all known cases of tailoring oxidation steps act on secondary carbon atoms (methylene groups). Oxidation of primary carbon atoms (methyl groups) in polyketide biosynthesis, as has presently been found, is unprecedented. No knowledge is available on the molecular basis of epoxide formation in polyketide products, though epoxides are present in a few known structures.

As has been mentioned above, each of the genes ORF1, ORF2 and ORF3 separately or in combination can be used for various purposes, as will be discussed in detail below.

Targeted inactivation of one or more of the present genes, e.g. through marker insertion or replacement with a non-functional gene equivalent, will interfere with at least one (oxidation) step in the Pimaricin biosynthetic route, and thus result in modified Pimaricin molecules characterized by a different oxidative state. E.g. molecules can be created lacking the epoxide function at carbons C4 and C5 , or molecules with a modified oxidation state of the carboxyl group at C12 resulting in an aldehyde, alcohol, or methyl group at this position.

Disruption of chromosomally encoded genes can be accomplished by gene replacement strategies: gene replacement is preferably carried out using suicide plasmid vectors or defective phage vectors carrying modified target genes and detection or selection marker genes. The various elements

useful for such strategies, and how to employ them, are described below.

Target gene modification can be accomplished by disruption of the coding sequence by insertion or deletion of nucleotides or nucleotide stretches. Such insertions or deletions may be of different size. Preferably, they are of a size of at least 2 nucleotides, excepting multiples of 3. Alternatively, the coding region of the target gene may be replaced by that of a (marker) gene conferring an easily detectable phenotype to cells transformed with this construct. Suitable examples of such replacement genes are lacZ, xylE, Green Fluorescent Protein, and genes for the biosynthesis of antibiotics, such as erythromycin, apramycin, hygromycin, and thiostrepton, and metabolite analogues, such as fluoroacetamide.

Transfer of the disrupted target gene to the Pimaricin production host resulting in *in vivo* gene inactivation can be accomplished by using e.g. suicide vector systems or a defective phage containing a fragment internal to the coding region of the target gene, or a variant of the gene inactivated through deletion or insertion of DNA stretches as described above, and in addition a detection or selection marker. Suicide vectors and defective phages are characterized by their inability to propagate autonomously in the strain to be transformed and thus cannot by themselves stably be maintained. For *Streptomycetes* in general several suicide systems are available: suicide vectors can be chosen from the group of extrachromosomal element based cloning vectors available for *E. coli*, which cannot replicate in *Streptomyces* species, like e.g. pBR322, pUC, CoID, RSF1010, RK2 and vectors derived from these plasmids. Similarly, *Streptomyces* plasmids characterized by a limited host range can be selected that are incapable of stable maintenance in the desired host strain. Examples of such narrow host range plasmids are SLP1.2 and SCP2, and cloning vectors derived from these plamids. Still another possibility is to use

temperature sensitive variants of *Streptomyces* wide host range plasmids. These plamids are characterized by their inability to replicate above a certain (restrictive) temperature. Besides non-replicative plasmids, also defective phage vectors have been developed based on the *Streptomyces* phage phiC31 and proven extremely useful for genetic analysis. In this regard, it is noted that an extensive overview of known *Streptomyces* genetic engineering techniques may be found in Hopwood et al. (D.A. Hopwood, M.J. Bibb, K.F. Chater, T. Kieser, C.J. Bruton, H.M. Kieser, D.J. Lydiate, C.P. Smith, J.M. Ward, H. Schrempf, Genetic Manipulation of *Streptomyces*: A Laboratory Manual, The John Innes Foundation, Norwich, England, 1985).

The above mentioned suicide constructs can be introduced in a desired host cell using transformation procedures with isolated DNA, by conjugation from a donor microorganism, e.g. an *E. coli* or *Streptomyces* strain harboring the construct, or via transfection by phage particles. All of these methods are well within the knowledge of the person skilled in the art.

Upon introduction of such a construct in the microorganism of interest, e.g. *Streptomyces natalensis*, stable maintenance of the introduced genetic information is only possible by integration of the construct in the host chromosome, preferably by homologous recombination with the chromosomal copy of the target gene. Strains having integrated the construct in the chromosome can be detected by the expression of the co-introduced marker: in case of a detection marker, transformed colonies can be screened for acquired properties like conversion of a colorless substrate into a colored compound (applicable with e.g. the genes lacZ, or xylE) or fluorescence (by expression of e.g. Green Fluorescent Protein). Alternatively, a marker can be used which allows selection of transformed strains by acquired resistance to e.g. antibiotics or toxic metabolite analogues; the latter method usually is employed more frequently because

only cells with the acquired resistance will be able to grow in media containing the antibiotic or toxic metabolite analogue. If an internal fragment of the target gene is used for the construction of the suicide vector or defective phage, integration of the construct into the chromosomal copy of the target gene will result in inactivation immediately. If the suicide construct or defective phage contains the complete target gene or a fragment including the N-terminal or C-terminal coding region, though inactivated through smaller insertions or deletions, only integration of the construct will result in the presence of an active and inactive copy of the gene, separated by vector DNA. For obtaining a strain with only an inactive copy, a second homologous recombination has to take place removing the vector sequences and the active copy of the target gene. Strains having undergone this second homologous recombination can be detected by the loss of the acquired property encoded by the co-introduced marker gene.

Another application of the present genes from the Pimaricin gene cluster lies in overexpression of one or more of these genes in the natural host (*Streptomyces natalensis*). The expression of the individual genes within the cluster is tightly regulated by the cell physiology and/or cluster specific regulatory genes. This internal control may be appropriate for production of the antibiotic in the natural environment, but undesirable for industrial production. Overexpression of all genes of the cluster by introduction of additional gene copies or altering the controlling elements (e.g. promoters or regulatory genes) can boost antibiotic production considerably, as has been shown for a.o. Actinorhodin production by *Streptomyces coelicolor*. A similar effect can be obtained by overexpression of specifically those genes encoding enzymes representing in rate limiting steps of the antibiotic biosynthesis.

Additional copies of each of the present genes from the Pimaricin biosynthesis gene cluster, either separately or

in different combinations, can be introduced into *Streptomyces natalensis* and increase the efficiency of the oxidative reactions leading to the natural Pimaricin molecule, thus resulting in strains displaying an improved Pimaricin production, expressed either in an increased Pimaricin titre in the culture broth or a higher product yield on substrate consumed. Of course, enhanced expression of certain genes can also be combined with inactivation of other genes, thus effecting improved production of variants of Pimaricin as described above.

Strains containing additional copies of target genes can be obtained through introduction of complete genes including expression signals (promoters) on the production host chromosome by techniques employing suicide vectors or defective phage as described above. Alternatively, autonomously replicating DNA molecules derived from phage genomes or extrachromosomal elements like plasmids can be used to carry the additional genes. Over the past two decades many naturally occurring *Streptomyces* plasmids have been turned into efficient cloning vectors, the most commonly used vectors being derived from plasmids pIJ101 and SCP2. Other vectors can be constructed based on the plasmid naturally occurring in *Streptomyces natalensis*, as disclosed in GB patent application nr 2210619 (Antibioticos), using selection and/or detection markers similar to those employed for the pIJ101 derived vectors, such as pIJ702, pIJ486, with or without added markers as described above.

For gene expression a large variety of promoters efficiently directing transcription of genes in Streptomyces is available; an example of a constitutive promoter is the ermE promoter, directing expression of the erythromycin resistance gene from *Saccharopolyspora erythraea*, whereas the agarase gene promoter from *S.coelicolor*, the promoter of the glycerol utilization operon, or the tipA promoter are examples of promoters inducible by specific substrates. Using techniques known in the art additional promoters can be

obtained e.g. promoters endogenous to *S.natalensis* (see J.M.Ward, G.R.Janssen, T.Kieser, M.J.Bibb, M.J.Buttner, M.J.Bibb. 1986. Mol.Gen.Genet. 203: 468-478).

The degree of overexpression can be manipulated by the choice of the promoter or the amount of inducing compound, and in addition by the choice of the autonomously replicating vector system; depending on the vector derivative used, predetermined plasmid copy numbers can range from 1 or 2 to 500. It is well within the expertise of the normal person skilled in the art to adjust the vector system to the desired degree of overexpression.

Both of the above uses of the present genes from the Pimaricin biosynthetic gene cluster, i.e. inactivation to obtain new variants of Pimaricin and overexpression to increase Pimaricin productivity, can also be applied to strains producing structurally similar bioactive compounds such as Amphotericin B (*Streptomyces nodosus*), Nystatin (*Streptomyces noursei*) (see Figure 4) etc.. to obtain variants of these compounds or improve productivity, or both. Using the present genes to inactivate the corresponding genes in other *Streptomyces* species will results in new derivatives of, *inter alia*, nystatin and amphotericin B altered in their oxidative state.

Still another application of the present genes is the heterologous expression and exploitation of the enzymatic activity encoded by one or more of these genes. Using similar vector systems as employed for overexpression of the oxidative genes in *S.natalensis*, other microorganisms, preferably *Streptomycetes* species, such as the easily accessible strain *Streptomyces lividans*, can be genetically transformed and thus acquire new oxidative enzymatic activity. This route is particularly useful for application of the enzymatic activities to the oxidative modification of other bioactive compounds, such as secondary metabolites, antibiotics, anticancer agents etc., which often are highly functionalized chemical entities. Thus, it is possible to

introduce one or more of the present genes in hosts producing such bioactive compounds naturally or have acquired the genetic information to produce compounds by recombinant DNA technology. Strains having acquired the genes encoding the oxidative enzymatic activity from the Pimarcin biosynthetic gene cluster will then be able to introduce epoxide functions or alcohol, aldehyde, or carboxyl groups into metabolites previously not modified in such a way. Thus, an approach has been provided, which allows for the creation of new variants of bioactive compounds not obtainable by chemical means (exemplified in Example 6 below).

The invention will now be elucidated by the following, non-restrictive examples.

EXAMPLES

Example 1. Isolation and identification of Pimaricin biosynthetic genes.

Streptomyces natalensis strain ATCC27448 was grown in YEME medium (D.A. Hopwood, M.J. Bibb, K.F. Chater, T. Kieser, C.J. Bruton, H.M. Kieser, D.J. Lydiate, C.P. Smith, J.M. Ward, H. Schrempf, Genetic Manipulation of *Streptomyces*: A Laboratory Manual, The John Innes Foundation, Norwich, England, 1985) at 30°C for 3 days. Mycelium was harvested and total DNA was extracted and purified essentially as described by Hopwood (ibid.).

Total *S.natalensis* DNA was subjected to partial digestion with restriction enzyme *Sau*3AI and size fractionated on 0.8% agarose gel. Fragments of 30-40 kbp were isolated, inserted into *Bam*HI digested cosmid Supercos1 and subsequently introduced in *E.coli* strain XL1-Blue MR according to protocols suggested by the supplier (Stratagene, La Jolla).

Thus a cosmid library of *S.natalensis* DNA in *E.coli* was obtained. This cosmid library was then screened for the

presence of polyketide synthase (PKS) related sequences by hybridization with radioactively labeled fragments from known PKS genes from the Rapamycin biosynthesis cluster from *Streptomyces hygroscopicus* (T.Schwecke, J.F.Aparicio, Y.Molnár, A.König, L.E.Khaw, S.F.Haydock, M.Oliynyk, P.Caffrey, J.Cortés, J.B.Lester, G.A.Böhm, J.Staunton, P.F.Leadlay. 1995. Proc. Natl. Acad. Sci. USA 92: 7839-7843).

Several clones were isolated as containing sequences hybridizing to a fragment containing the KS module 5 of of rapB.

Complete DNA sequence determination of a number of neighbouring *Not*I fragments from Cos9 was performed after cloning the fragments in pBluescript. Computer assisted analysis of the DNA sequences revealed the presence of genes clearly identifiable as PKS gene modules by nucleotide and derived amino acid sequence homology with established PKS genes and proteins, involved in the biosynthesis of erythromycin and rapamycin, as well as with fatty acid synthase genes and proteins, which catalyze a similar set of reactions. The complete nucleotide sequences and derived amino acid sequences of two Pimaricin PKS genes are given as SEQ ID numbers 1-4. Using a 0.7 kb *Not*I fragment from Cos9 as a probe, the extent of the PKS related genes on the cosmid map was established as indicated in Figure 1.

Example 2. PKS genes are essential for Pimaricin biosynthesis

A completely sequenced 3.3 kb *Not*I DNA fragment (see Figure 1)(in pBluescript), encoding (part of) a *S.natalensis* PKS as deduced form the organizational and structural sequence similarities with known PKS, was excised by *Sac*I from the sequencing vector, subcloned in the phage vector KC515 (M.R.Rodicio, C.J.Bruton, K.F.Chater. 1985. Gene 34: 283-292) and introduced in *S.lividans* to obtain infectious

particles (recombinant phage) containing the *S.natalensis* PKS fragment. Infection of *S.natalensis* using this recombinant phage population and selection for resistance to the antibiotic viomycin, allowed the isolation of lysogens, originated through integration of the recombinant phage DNA into the *S.natalensis* chromosomal DNA by homologous recombination of the PKS regions.

None of 20 lysogens tested displayed antifungal activity as analyzed by an agar plate bioassay using *Candida utilis* as the indicator organism. Detailed analysis of one of the lysogens by Southern hybridization studies confirmed that indeed integration of the recombinant phage DNA into the *S.natalensis* chromosomal PKS locus had occurred.

Culturing the lysogen with the disrupted PKS gene in standard production medium (25 g/l soya peptone, 0.5 mM $ZnSO_4$, 20 g/l glucose, pH 7.5) followed by extraction of the culture broth with butanol, and UV spectrophotometric analysis indicated that no traces of Pimaricin were produced by this lysogen (J.F.Martín, A.L.Demain. 1975. Biochem. Biophys. Res. Commun. 71: 1103-1109).

Example 3. Detailed sequence analysis of non-PKS genes; preliminary identification.

Full sequence analysis of the regions flanking the PKS genes of Example 1 revealed the presence of additional open reading frames (ORF) potentially encoding proteins functional in Pimaricin biosynthesis.

Homology comparison of the deduced amino acids sequences of the ORFs indicated the involvement of several in oxidation/reduction reactions: ORF1 showed a clear homology with previously identified cholesterol oxidases, ORF2 and ORF3 were similar to cytochrome P-450 monooxygenase proteins. Also genes encoding accessory proteins for the P-450 enzymes seem to be present i.e. ferredoxin type. Complete nucleotide

sequences of the respective genes and derived amino acid sequences are added as SEQ ID numbers 5-10.

Example 4. Functional characterization of Pimaricin biosynthesis non-PKS genes.

To define the involvement of the accessory genes/proteins in Pimaricin biosynthesis, the function each of the major ORFs, i.e. ORF1, ORF2, and ORF3, was disrupted and the effect on Pimaricin production established. Similar strategies as described in Example 2 for the PKS disruption were employed for the non-PKS genes. Detailed information on the chromosomal regions enompassing the three open reading frames (ORF's) is presented in Figure 2.

ORF1: a 7kb *Sph*I fragment containing the complete ORF1 was cloned into pUC19, the resulting plasmid was digested with *Bgl*II, the cohesive ends were filled in by treatment with Klenow polymerase and religated. This new plasmid was used as a source for DNA for the gene replacement. The 2.9 kb *Bam*HI-*Pst*I fragment from the plasmid was cloned into the *Bam*HI-*Pst*I siteds of KC515. Lysogens were obtained by selection for thiostrepton. The second recombination event was searched for by the loss of thiostrepton resistance. The insertion and subsequent loss of the phage was confirmed by Southern hybridization.

ORF2: a 1.6 kb *Sal*I fragment encompassing most of ORF2 was inserted in pUC19; the resulting plasmid was cut with restriction enzyme *Bst*BI and treated with T4 DNA polymerase to disrupt the presumed coding sequence, followed by religation. Again a gene replacement strategy was followed required in which the intact chromosomal copy of ORF2 was exchanged for the disrupted variant. Thus the *Bam*HI-*Pst*I fragment of the plasmid was transferred to the phage vector KC515, via which lysogens were obtained as described above, having the recombinant phage DNA inserted in the

corresponding *S.natalensis* chromosomal locus. For disruption of the gene a second homologous recombination event was selected for by loss of thiostrepton resistance. Thiostrepton sensitive strains derived from the ORF2 lysogens were analyzed for the presence of the intact or disrupted copy by Southern hybridization.

Alternatively, a 1.95 kb *Bam*HI fragment can be used after cloning in pUC19/*Bam*HI, and transfer, after modification of the *Bst*BI restriction site, of the *Bam*HI fragment to phage vector KC515.

ORF3: direct disruption was accomplished by insertion of a 987 bp *Pvu*II fragment internal to ORF3 into the phage vector KC515 and transformation of *S.lividans* to obtain infectious recombinant phage particles. After transfection of *S.natalensis* lysogens were isolated as described above. In a different approach gene disruption is realized by insertion of a 3.4 kb SalI fragment encompassing the complete ORF3 in *Sal*I digested pUC19, followed by digestion of the resulting plasmid with *Mlu*I and by T4 DNA polymerase treatment, and religation. Subsequently, the *Bam*HI-*Pst*I fragment of the plasmid was transferred to the phage vector KC515, via which lysogens were obtained as described above, having the recombinant phage DNA inserted in the corresponding *S.natalensis* chromosomal locus. For disruption of the gene a second homologous recombination event was selected for by loss of thiostrepton resistance. Thiostrepton sensitive strains derived from the ORF3 lysogens were analyzed for the presence of the intact or disrupted copy by Southern hybridization.

Example 5. Analysis of ORF1, ORF2, and ORF3 gene disruptants of *S.natalensis*.

Each of the separate gene disruptants was analyzed using the bioassay with *C.utilis*. None displayed an

antifungal activity as distinct as the wild-type strain *S.natalensis* ATCC27448.

All disruptions combined in one strain also did not produce *in vivo* antifungal activity. In the latter strain upon growth in pimaricin production medium (see Example 2) an analysis of total metabolites from the culture broth indicated the presence of a Pimaricin precursor, wherein C4, C5 and the carbon atom of the carboxyl group at C12 are unoxidized carbon atoms at (see Figure 3B). This was determined after extraction of the whole broth with methanol using NMR spectroscopy.

Example 6. Overexpression of ORF1, ORF2, and ORF3 in *S.natalensis*.

Overexpression of ORF1, ORF2 and ORF3 separately was obtained by placing each gene under the direction of the ermE promoter from *Saccharopolyspora erythraea* (M.J. Bibb, G.R. Janssen, J.M. Ward. 1985. Gene 38: 215-226). A useful derivative of this promoter, having a number of cloning sites attached was obtained by PCR using the following oligonucleotides: SEQ ID 11:
AAACTGCAGCTCTAGAGGCGGCTTGCGCCCGATGCTAGTC

SEQ ID 12:
AAACTGCAGCTCTAGATGCCCGGGTATCGATCGTCGACGGCATGCGGATCCTACCAACCGG CACGATTG

The 225 bp PCR fragment obtained was digested with *Pst*I, purified by agarose gel electrophoresis and inserted in PstI digested pUC19, yielding pUCermE

ORF1 was inserted in pUCermE as a 2.2 kb *Sph*I-*Cla*I fragment encompassing the complete coding sequence, for ORF2 the 3.5 kb *Cla*I-*Nru*I fragment was used, and for ORF3 the 2.8 kb *Sal*I-*Kpn*I fragment. Each ermE promoter-ORF combination was

subsequently excised as a *Pst*I fragment, inserted in *Pst*I digested phage vector KC515 and introduced in S.natalensis essentially as described in Example 4.

Recombinant *S.natalensis* strains thus obtained, overexpressing each of the three genes separately, showed improved Pimaricin production levels by 10 -15 % after growth under standard production conditions (see Example 2).

Example 7. Functional expression of *S.natalensis* ORF1, ORF2, and ORF3 in *S.lividans*.

The ermE promoter-ORF combinations were excised from the corresponding pUCermE construct as an *Xba*I fragment, inserted in *Xba*I digested plasmid pIJ486 (J.M.Ward, G.R.Janssen, T.Kieser, M.J.Bibb, M.J.Buttner, M.J.Bibb. 1986. Mol.Gen.Genet. 203: 468-478) and introduced by transformation into *S.lividans*. Enzymatic assays confirmed the activity of the different ORFs as oxidizing enzymes. Surprisingly, in the presence of artificial oligoketides synthesized by assembled PKS modules from the Erythromycin pathway, oxidizing activity on the exo-methyl groups of these oligoketides was observed similar to the oxidations observed in the Pimaricin polyketide.

SEQUENCE LISTING

<110> DSM N.V.

<120> Genes encoding enzymes in the biosynthetis of pimaricin and applications thereof

<130> EP-2959 P

<140>
<141>

<160> 12

<170> PatentIn Ver. 2.1

<210> 1
<211> 5544
<212> DNA
<213> Streptomyces natalensis

<220>
<221> CDS
<222> (1)..(5544)

<400> 1

```
atg gtg ccc gtc cac aca gat gac tac gcg atc cag ccg ccc gcc gac   48
Met Val Pro Val His Thr Asp Asp Tyr Ala Ile Gln Pro Pro Ala Asp
  1               5                  10                  15

acc gcg cat gga ggg ggt ggc ttc aca ctg cct gct gtg ttc gag gcc   96
Thr Ala His Gly Gly Gly Gly Phe Thr Leu Pro Ala Val Phe Glu Ala
             20                  25                  30

gcc gtg gaa tcc gcc ccc gac gca gtg gcc ctc gtc gac ggc acc gtt  144
Ala Val Glu Ser Ala Pro Asp Ala Val Ala Leu Val Asp Gly Thr Val
         35                  40                  45

cct gga cct ggg cgc atg tgg cgc gca gac gtc gat gcg ctg gcc cgc  192
Pro Gly Pro Gly Arg Met Trp Arg Ala Asp Val Asp Ala Leu Ala Arg
     50                  55                  60

ggg ttg cag gag tcg ggt atc gca ccc ggt gac gtg gtg gcg gtg cgg  240
Gly Leu Gln Glu Ser Gly Ile Ala Pro Gly Asp Val Val Ala Val Arg
 65                  70                  75                  80

ctg ccg aac tgc ggg agg ttt ccc acc ctg cac ctg gcc gtc gcg gcc  288
Leu Pro Asn Cys Gly Arg Phe Pro Thr Leu His Leu Ala Val Ala Ala
                 85                  90                  95

gtc gga gcg gta ctg ctg ccc atc cac cag ggc acc ccg ctc ccg gag  336
Val Gly Ala Val Leu Leu Pro Ile His Gln Gly Thr Pro Leu Pro Glu
            100                 105                 110

gtc gac gcg ctg ctg acc cgg gcg gaa ccc gcc ctt ctc gtc ctg tcg  384
Val Asp Ala Leu Leu Thr Arg Ala Glu Pro Ala Leu Leu Val Leu Ser
        115                 120                 125

gcc gcc ggg agc gac ggc ctg gcg acg gcc cgt tcg ctt ctg gag agc  432
Ala Ala Gly Ser Asp Gly Leu Ala Thr Ala Arg Ser Leu Leu Glu Ser
    130                 135                 140

gtg ccg tcg ctg cgc ggc gtg ctg ctg gcc ggg gcg tcg gga gac ggg  480
Val Pro Ser Leu Arg Gly Val Leu Leu Ala Gly Ala Ser Gly Asp Gly
145                 150                 155                 160
```

```
gag tcc gga tcc gtc gga ggt ggg gag tca gga tcc ggg cgt cgg tcg   528
Glu Ser Gly Ser Val Gly Gly Gly Glu Ser Gly Ser Gly Arg Arg Ser
                165                 170                 175

ctg gac ggg ctg ctg gcc ggc tgg gcg ggc agc ggg ccc cgg ccg gtg   576
Leu Asp Gly Leu Leu Ala Gly Trp Ala Gly Ser Gly Pro Arg Pro Val
            180                 185                 190

gac gtc acc ccc gat atg ccg ctc gtt ctc gtc ccc tcg tcc ggg acg   624
Asp Val Thr Pro Asp Met Pro Leu Val Leu Val Pro Ser Ser Gly Thr
        195                 200                 205

gtc tcg gca cgg ccc aaa ctg tgt gtg cac agc cac gac ggt ctg ctg   672
Val Ser Ala Arg Pro Lys Leu Cys Val His Ser His Asp Gly Leu Leu
    210                 215                 220

tcg aac acc gca gcc gtc acc gcc gag gcc gcc gac gcc ttc gac ggt   720
Ser Asn Thr Ala Ala Val Thr Ala Glu Ala Ala Asp Ala Phe Asp Gly
225                 230                 235                 240

ccg gtc ctc acg gcc tgc ccg atg acc cat ctg ttc ggc ctc cag tcc   768
Pro Val Leu Thr Ala Cys Pro Met Thr His Leu Phe Gly Leu Gln Ser
                245                 250                 255

ctg cac gcg gcg ctg ttc gcc gcc tgc acg cag gtc ctg ctc acc ggc   816
Leu His Ala Ala Leu Phe Ala Ala Cys Thr Gln Val Leu Leu Thr Gly
            260                 265                 270

tgg gac gtg gac cgc ttc ctg gaa cag gct cgg gag cac ggc ccc cgt   864
Trp Asp Val Asp Arg Phe Leu Glu Gln Ala Arg Glu His Gly Pro Arg
        275                 280                 285

gtc gtg ttc gcc gtg ccc gcc caa ctg cgg gac gtc gtg aca cga ctg   912
Val Val Phe Ala Val Pro Ala Gln Leu Arg Asp Val Val Thr Arg Leu
    290                 295                 300

gcc agg acc ggc gaa ccg gcg ggc ttc acg ccg tac cag gtg cgt acc   960
Ala Arg Thr Gly Glu Pro Ala Gly Phe Thr Pro Tyr Gln Val Arg Thr
305                 310                 315                 320

gcg ggg gcc gcc gtc gcg ccc gcg ctc gcc gta cgg gta cgc gcc gtc  1008
Ala Gly Ala Ala Val Ala Pro Ala Leu Ala Val Arg Val Arg Ala Val
                325                 330                 335

ctc gac tgc gaa ctg gtc gtg gtg tgg ggg atg tcc gag atc ggc acc  1056
Leu Asp Cys Glu Leu Val Val Val Trp Gly Met Ser Glu Ile Gly Thr
            340                 345                 350

ggc acc cgc acc cgg gcc cac cac ccc gac ggc tgt gtg gga gag ccg  1104
Gly Thr Arg Thr Arg Ala His His Pro Asp Gly Cys Val Gly Glu Pro
        355                 360                 365

gtc agc ggg gtg gac gtg cgc gtc gtc gac gag cac ggc caa gag tgc  1152
Val Ser Gly Val Asp Val Arg Val Val Asp Glu His Gly Gln Glu Cys
    370                 375                 380

gcg gcg gac gag agg ggt gaa ctc cag tac cgg ggg cca ggg ttg ttc  1200
Ala Ala Asp Glu Arg Gly Glu Leu Gln Tyr Arg Gly Pro Gly Leu Phe
385                 390                 395                 400

cgc ggc tat ttc cgc gaa ccg gag ctg acg cgc tcg gcc ctc acc gac  1248
Arg Gly Tyr Phe Arg Glu Pro Glu Leu Thr Arg Ser Ala Leu Thr Asp
                405                 410                 415
```

```
gac ggc tgg ctg cgg acc ggc gat ctc gcg acc gtc gac gcg gac ggc   1296
Asp Gly Trp Leu Arg Thr Gly Asp Leu Ala Thr Val Asp Ala Asp Gly
            420                 425                 430

gtg gtg gtg ctg cac ggc cgg gcg gcc gag ctg atc aat acg ggg ggc   1344
Val Val Val Leu His Gly Arg Ala Ala Glu Leu Ile Asn Thr Gly Gly
        435                 440                 445

cgg aag ttc agc gcc ggc gag gtc gag gga ctg ctc tcc ggc ttt acg   1392
Arg Lys Phe Ser Ala Gly Glu Val Glu Gly Leu Leu Ser Gly Phe Thr
    450                 455                 460

gac ctg ggc ccg ctg gcc gtc gtc ggc gca ccg gac gac cgg ctc ggg   1440
Asp Leu Gly Pro Leu Ala Val Val Gly Ala Pro Asp Asp Arg Leu Gly
465                 470                 475                 480

gag tac ccg tgc ctg gtc gtg acc gac cac gcg gac ggc acc atc ggc   1488
Glu Tyr Pro Cys Leu Val Val Thr Asp His Ala Asp Gly Thr Ile Gly
                485                 490                 495

ctg agc gag gtg acc gcg ttc ctg cgc cgg ctg ggg ctc gcc gac cac   1536
Leu Ser Glu Val Thr Ala Phe Leu Arg Arg Leu Gly Leu Ala Asp His
            500                 505                 510

aag atc ccc ctc gaa ctg gtc acc gtg cgc gag ctt ccc ttc tcc cct   1584
Lys Ile Pro Leu Glu Leu Val Thr Val Arg Glu Leu Pro Phe Ser Pro
        515                 520                 525

gcc ggg aag ctc gac cgc ggg gcg ctc aag cgg ctg ctc gcc aac ctt   1632
Ala Gly Lys Leu Asp Arg Gly Ala Leu Lys Arg Leu Leu Ala Asn Leu
    530                 535                 540

gcg gag gtg tcc gtt ccg gcc cgt ctc ggc gcg gtc ccg ccc tac acg   1680
Ala Glu Val Ser Val Pro Ala Arg Leu Gly Ala Val Pro Pro Tyr Thr
545                 550                 555                 560

gcc gag gag gcg ctg gac ctc gta cgc gac tgc gtc ggc cgg gtg ctc   1728
Ala Glu Glu Ala Leu Asp Leu Val Arg Asp Cys Val Gly Arg Val Leu
                565                 570                 575

cgt tac ggc ggc gcg gcc gtc ccc ttc ccc ccg gac aag gac ttc ttc   1776
Arg Tyr Gly Gly Ala Ala Val Pro Phe Pro Pro Asp Lys Asp Phe Phe
            580                 585                 590

tcc ccg gac aag gac ttc cgc cag ctc ggc ctg gac tcc atc ggc gcg   1824
Ser Pro Asp Lys Asp Phe Arg Gln Leu Gly Leu Asp Ser Ile Gly Ala
        595                 600                 605

gtg cgg ctg cgg aat ctg ctg cgg gag gag acc ggg ctg ccg ctc ccg   1872
Val Arg Leu Arg Asn Leu Leu Arg Glu Glu Thr Gly Leu Pro Leu Pro
    610                 615                 620

gcc acc ctg gcc ttc gac tcc ccc act ccg cga gcc gtc gcg cgc gtc   1920
Ala Thr Leu Ala Phe Asp Ser Pro Thr Pro Arg Ala Val Ala Arg Val
625                 630                 635                 640

ctg gcc gag cag gag gag ccg tcc cag gac gag ccg agg gag aac ccg   1968
Leu Ala Glu Gln Glu Glu Pro Ser Gln Asp Glu Pro Arg Glu Asn Pro
                645                 650                 655

gcg gac ggt gcc gac ccg gtg gcg atc gtg ggc atg gcc tgc cgg ctg   2016
Ala Asp Gly Ala Asp Pro Val Ala Ile Val Gly Met Ala Cys Arg Leu
            660                 665                 670

ccg ggc gga gcc gac tcc ccc gac gcc ctg tgg gag ctg ctc gcc gac   2064
```

```
Pro Gly Gly Ala Asp Ser Pro Asp Ala Leu Trp Glu Leu Leu Ala Asp
        675                 680                 685

ggg acc gac gcg atg tcc ccc ttc ccc acg gac cgc ggc tgg gac ctg   2112
Gly Thr Asp Ala Met Ser Pro Phe Pro Thr Asp Arg Gly Trp Asp Leu
    690                 695                 700

gac cgg ctg ttc gac gag gat gcc gac cgc ccg ggt acc tcg tac gcc   2160
Asp Arg Leu Phe Asp Glu Asp Ala Asp Arg Pro Gly Thr Ser Tyr Ala
705                 710                 715                 720

cgc gaa ggc ggc ttc ctg cac gac gcg ggc gac ttc gac gcg ggc ttc   2208
Arg Glu Gly Gly Phe Leu His Asp Ala Gly Asp Phe Asp Ala Gly Phe
                725                 730                 735

ttc ggc ctg tcg gac cag gag gcg acg gcg acc gat ccg cag cag cgg   2256
Phe Gly Leu Ser Asp Gln Glu Ala Thr Ala Thr Asp Pro Gln Gln Arg
            740                 745                 750

ctg ctt ctg gag gcg gcc tgg gag acc ttc gag cgg gcg ggc atc gac   2304
Leu Leu Leu Glu Ala Ala Trp Glu Thr Phe Glu Arg Ala Gly Ile Asp
        755                 760                 765

ccg cag tcc ctg agg gga agc cgt acg ggc gtg ttc acg ggc gcg atg   2352
Pro Gln Ser Leu Arg Gly Ser Arg Thr Gly Val Phe Thr Gly Ala Met
    770                 775                 780

gac cgc ggc tat gga acc agc gcg tcc gcc gcg ccc agc gca tgg gag   2400
Asp Arg Gly Tyr Gly Thr Ser Ala Ser Ala Ala Pro Ser Ala Trp Glu
785                 790                 795                 800

agc atg ctc atc acc ggg acc gcc ggc agc gcg gtc tcg ggg cgc atc   2448
Ser Met Leu Ile Thr Gly Thr Ala Gly Ser Ala Val Ser Gly Arg Ile
                805                 810                 815

gcc tac acc tac ggg ctc gaa ggc ccc gcg ctg acg gtc gac acc gcc   2496
Ala Tyr Thr Tyr Gly Leu Glu Gly Pro Ala Leu Thr Val Asp Thr Ala
            820                 825                 830

tcc tcg tcc tcc ctc gtc gcc ctg cat ctg gcc tgc cgg tcg ctg cgc   2544
Ser Ser Ser Ser Leu Val Ala Leu His Leu Ala Cys Arg Ser Leu Arg
        835                 840                 845

tcg ggc gag acc gac ctg gcg ctg gcc ggc ggc gtc acc gtc atg gcg   2592
Ser Gly Glu Thr Asp Leu Ala Leu Ala Gly Gly Val Thr Val Met Ala
    850                 855                 860

acc ccg gcg ccc ttc gca cac ttc tcc cgg ctg cgc gcg ctg tcc ccc   2640
Thr Pro Ala Pro Phe Ala His Phe Ser Arg Leu Arg Ala Leu Ser Pro
865                 870                 875                 880

gac tcc cgc tcc atg gcg tac gcg gac gcc gcg aac ggc tcg gcg tgg   2688
Asp Ser Arg Ser Met Ala Tyr Ala Asp Ala Ala Asn Gly Ser Ala Trp
                885                 890                 895

tcg gag ggc gcg ggg ctg ctt ctg ctg gag cgg ctg agc gac gcc cgg   2736
Ser Glu Gly Ala Gly Leu Leu Leu Leu Glu Arg Leu Ser Asp Ala Arg
            900                 905                 910

cgc aac gga cac cgt gtc ctg gcg ctc gta cgg ggc tcc gcc gtg aat   2784
Arg Asn Gly His Arg Val Leu Ala Leu Val Arg Gly Ser Ala Val Asn
        915                 920                 925

cag gac ggc gcc tcc aac ggg ctc acc gcg ccg agc gga ccc gca cag   2832
Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Ser Gly Pro Ala Gln
```

```
    930                 935                 940

cag cgc gtc atc cgc cag gcc ctg gcc gac gcc ggg ctg acg ccg cag   2880
Gln Arg Val Ile Arg Gln Ala Leu Ala Asp Ala Gly Leu Thr Pro Gln
945                 950                 955                 960

gac gtg gac gcc gtg gag ggg cac ggc acc ggc acg ccg ctc ggc gac   2928
Asp Val Asp Ala Val Glu Gly His Gly Thr Gly Thr Pro Leu Gly Asp
                965                 970                 975

ccc atc gag gcg cag gcg ctg ctg gcc acg tac ggc caa cag cgg cct   2976
Pro Ile Glu Ala Gln Ala Leu Leu Ala Thr Tyr Gly Gln Gln Arg Pro
            980                 985                 990

gtg gaa cgg ccg ttg tgg ctg ggg tcg gtg aag tcg aac ttc ggg cac   3024
Val Glu Arg Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Phe Gly His
        995                1000                1005

aca caa gcc gcc gcc ggg gtc gtc ggc gtc atc aag acg gtg ctc gcg   3072
Thr Gln Ala Ala Ala Gly Val Val Gly Val Ile Lys Thr Val Leu Ala
   1010                1015                1020

ctg cgc cac ggc gtc ctg ccg cag acg ctg cac gtg gac gct ccc tcg   3120
Leu Arg His Gly Val Leu Pro Gln Thr Leu His Val Asp Ala Pro Ser
1025                1030                1035                1040

gcc aag gtg gac tgg tcc gcc ggt tcg gta cgg ctg ctg acc gag gcg   3168
Ala Lys Val Asp Trp Ser Ala Gly Ser Val Arg Leu Leu Thr Glu Ala
                1045                1050                1055

cgg ccc tgg cca cgg gag agc gga cgt acg cgc cgg gcg ggg gtg tcc   3216
Arg Pro Trp Pro Arg Glu Ser Gly Arg Thr Arg Arg Ala Gly Val Ser
            1060                1065                1070

tcg ttc ggg ctc acc ggc acc aac gcg cac gtg atc ctg gag gag gcg   3264
Ser Phe Gly Leu Thr Gly Thr Asn Ala His Val Ile Leu Glu Glu Ala
        1075                1080                1085

ccg gga gag gcg gcg gca ggg gcg cgg gcc gag gtt ccc gag gag gcg   3312
Pro Gly Glu Ala Ala Ala Gly Ala Arg Ala Glu Val Pro Glu Glu Ala
   1090                1095                1100

cgg tgc gcc tcc tca ccg gct cga ctc ccg gag ccg ccc ggc gac gcg   3360
Arg Cys Ala Ser Ser Pro Ala Arg Leu Pro Glu Pro Pro Gly Asp Ala
1105                1110                1115                1120

gcc gcg ccc tgg gtg ctg tcc gcc cgg agc cgg gcg gcg ctg cgc gcg   3408
Ala Ala Pro Trp Val Leu Ser Ala Arg Ser Arg Ala Ala Leu Arg Ala
                1125                1130                1135

cag gcg ctc cgc ctg gcc gac cag gtg gcc gcc gac ccc ggt cta cgg   3456
Gln Ala Leu Arg Leu Ala Asp Gln Val Ala Ala Asp Pro Gly Leu Arg
            1140                1145                1150

gcc cag gat gtc gcc cat gcc ctg gcc acc tcc cgt acc ctg cac cgg   3504
Ala Gln Asp Val Ala His Ala Leu Ala Thr Ser Arg Thr Leu His Arg
        1155                1160                1165

cac cgg gcc gtc gtc agc ggc tcc gac cgg gca caa atg ctc gcc gcg   3552
His Arg Ala Val Val Ser Gly Ser Asp Arg Ala Gln Met Leu Ala Ala
   1170                1175                1180

gca aag cgg ttc ggg ctc ggt gag cgg acc gcg ggc gtc acc ccg gac   3600
Ala Lys Arg Phe Gly Leu Gly Glu Arg Thr Ala Gly Val Thr Pro Asp
1185                1190                1195                1200
```

```
gat tcc gcg ccg ggc ctg ctg gcc ttc gtc ttc tcc ggg cag ggc agc   3648
Asp Ser Ala Pro Gly Leu Leu Ala Phe Val Phe Ser Gly Gln Gly Ser
                1205                1210                1215

cag cgc agc ggc atg ggg cgc gcg gcg gcc gag gcg ttc ccg gtc ttc   3696
Gln Arg Ser Gly Met Gly Arg Ala Ala Ala Glu Ala Phe Pro Val Phe
            1220                1225                1230

gga cgg gcg ctg ggc gag gtg tgc gcc gcg ctg gac ccg ctg ctg aca   3744
Gly Arg Ala Leu Gly Glu Val Cys Ala Ala Leu Asp Pro Leu Leu Thr
        1235                1240                1245

cgc cca ctg acc tcg gtg atg tgg gcg gct ccc ggc tcc gag gag gcg   3792
Arg Pro Leu Thr Ser Val Met Trp Ala Ala Pro Gly Ser Glu Glu Ala
    1250                1255                1260

gcc cgt ctc gac gac acc acc tac acg cag ccc gcc ctg ttc gcc gtc   3840
Ala Arg Leu Asp Asp Thr Thr Tyr Thr Gln Pro Ala Leu Phe Ala Val
1265                1270                1275                1280

cag gtc gcc ctg tac cgg ctg ttc gag tcc tgg ggc gtg gtg ccg gac   3888
Gln Val Ala Leu Tyr Arg Leu Phe Glu Ser Trp Gly Val Val Pro Asp
                1285                1290                1295

cag ctg gtg ggg cat tcg gtc ggc gag atc tcc gcc gcc cat gtg gca   3936
Gln Leu Val Gly His Ser Val Gly Glu Ile Ser Ala Ala His Val Ala
            1300                1305                1310

ggc gtg ctc ggc ctc cgg gac gcg tgc acc ctg gtg gcg gcc cgt agc   3984
Gly Val Leu Gly Leu Arg Asp Ala Cys Thr Leu Val Ala Ala Arg Ser
        1315                1320                1325

agg ctg atg ggc gcg ctg ccg ccc ggc ggt gcg atg gtg gcg gta cgc   4032
Arg Leu Met Gly Ala Leu Pro Pro Gly Gly Ala Met Val Ala Val Arg
    1330                1335                1340

atc acg gaa ccc gaa gtg acc cca tgg ctc gcg gag ttg acg gac gag   4080
Ile Thr Glu Pro Glu Val Thr Pro Trp Leu Ala Glu Leu Thr Asp Glu
1345                1350                1355                1360

gtg tcg atc gcg gcc gtc aac ggt ccg cac tcc ctc gtg ctc gcg ggc   4128
Val Ser Ile Ala Ala Val Asn Gly Pro His Ser Leu Val Leu Ala Gly
                1365                1370                1375

gcc gag gcc ccg ctc gtc gcc ctc acg gac cgg ctc gcc gcc gcc gga   4176
Ala Glu Ala Pro Leu Val Ala Leu Thr Asp Arg Leu Ala Ala Ala Gly
            1380                1385                1390

cac aag acc cgg cgc ctc atg gtg agc acc gcg ccc cac tcg ccg ctg   4224
His Lys Thr Arg Arg Leu Met Val Ser Thr Ala Pro His Ser Pro Leu
        1395                1400                1405

atg gac ccc atg ctg gag gag ttc cgc gcg gtc gtc cgc acg ctg tcc   4272
Met Asp Pro Met Leu Glu Glu Phe Arg Ala Val Val Arg Thr Leu Ser
    1410                1415                1420

tac gcc gcg ccc gcc gtt ccc ctc gtc tcc acc gtc acc ggc cgc ccg   4320
Tyr Ala Ala Pro Ala Val Pro Leu Val Ser Thr Val Thr Gly Arg Pro
1425                1430                1435                1440

ctg acc ggc gag gag gcg cgc gac ccg gac cac tgg gtg cgg cat gtg   4368
Leu Thr Gly Glu Glu Ala Arg Asp Pro Asp His Trp Val Arg His Val
                1445                1450                1455
```

```
cgg cag tcc gtc cgc ttc aag gac gcg atc ggc cgg ctc cgg gac gaa   4416
Arg Gln Ser Val Arg Phe Lys Asp Ala Ile Gly Arg Leu Arg Asp Glu
            1460                1465                1470

cgc gtc acc ggg ttc ctg gag ctg ggt gcc gaa ccg gca ctc aca ccg   4464
Arg Val Thr Gly Phe Leu Glu Leu Gly Ala Glu Pro Ala Leu Thr Pro
        1475                1480                1485

atg atc gac gag tgc ctg gag tcc gcc gac ggg cag ccc ggg acc gcc   4512
Met Ile Asp Glu Cys Leu Glu Ser Ala Asp Gly Gln Pro Gly Thr Ala
    1490                1495                1500

ctg gtg ccg agt ctg cgc gcc gga gtg ccg gag cgg gat gcc ctg ctc   4560
Leu Val Pro Ser Leu Arg Ala Gly Val Pro Glu Arg Asp Ala Leu Leu
1505                1510                1515                1520

acc gcg gtc gcc cgg gtg cac gcc cag ggc gtt ccc gtc gac tgg gac   4608
Thr Ala Val Ala Arg Val His Ala Gln Gly Val Pro Val Asp Trp Asp
                1525                1530                1535

gcg gtg ctc ccc ggg gcc gag gct tcc gtc acc gtg cgc gga ctg ccc   4656
Ala Val Leu Pro Gly Ala Glu Ala Ser Val Thr Val Arg Gly Leu Pro
            1540                1545                1550

gcc gcc gac cgc cag tgg ttc cgc ttc gtc ccc gac cag ggc gcg ccg   4704
Ala Ala Asp Arg Gln Trp Phe Arg Phe Val Pro Asp Gln Gly Ala Pro
        1555                1560                1565

ctc acc ctc gcc gac cgc tcg ctg cac ctg gag ggc gcc gcc cac ctc   4752
Leu Thr Leu Ala Asp Arg Ser Leu His Leu Glu Gly Ala Ala His Leu
    1570                1575                1580

cgc gac gtg ggc ggc tgt cgc acc gcc gac ggc cgg tgg gtg aaa atg   4800
Arg Asp Val Gly Gly Cys Arg Thr Ala Asp Gly Arg Trp Val Lys Met
1585                1590                1595                1600

ggc gtc ctg tac cgc acc aac aat ctg cac gcg ctc acc gac gcc gac   4848
Gly Val Leu Tyr Arg Thr Asn Asn Leu His Ala Leu Thr Asp Ala Asp
                1605                1610                1615

ctg gcc aag ctc cag cgc ctc ggc atc cgt acg gac ttc gac ctg cgc   4896
Leu Ala Lys Leu Gln Arg Leu Gly Ile Arg Thr Asp Phe Asp Leu Arg
            1620                1625                1630

atg ccg ggc gag cgc gcc aaa gcc ccg aac cgc gtg ccc acg ggc gcg   4944
Met Pro Gly Glu Arg Ala Lys Ala Pro Asn Arg Val Pro Thr Gly Ala
        1635                1640                1645

cgc tat atc gtc gcc gat gcc ttt gac gca cat ctg cgg gac ggg ctg   4992
Arg Tyr Ile Val Ala Asp Ala Phe Asp Ala His Leu Arg Asp Gly Leu
    1650                1655                1660

agt ctg agc gag cgg gac ctg gag cgg ctg cgc gag gcc cgg ctg tcc   5040
Ser Leu Ser Glu Arg Asp Leu Glu Arg Leu Arg Glu Ala Arg Leu Ser
1665                1670                1675                1680

ggt ctc gac gag gcc gaa cag cac gcg ctc gta ctg gcc ctg gtc ctc   5088
Gly Leu Asp Glu Ala Glu Gln His Ala Leu Val Leu Ala Leu Val Leu
                1685                1690                1695

gcc gag acc tcg gcc gtg ctc ggc ggc cag gag acg cct ggc gag gag   5136
Ala Glu Thr Ser Ala Val Leu Gly Gly Gln Glu Thr Pro Gly Glu Glu
            1700                1705                1710

ccg cac ggc gag gag ggc cac cgc acg ttc aag gag atg ggc atc aac   5184
```

```
Pro His Gly Glu Glu Gly His Arg Thr Phe Lys Glu Met Gly Ile Asn
       1715                1720                1725

tcg ttg aac gcc gtc gaa ctg cgc aac cgc ctt atc gca gcc acg gac   5232
Ser Leu Asn Ala Val Glu Leu Arg Asn Arg Leu Ile Ala Ala Thr Asp
   1730                1735                1740

ctg cgg ctt ccc gcc acg ctc gtc tac gac tac ccc acg ccg aag gcc   5280
Leu Arg Leu Pro Ala Thr Leu Val Tyr Asp Tyr Pro Thr Pro Lys Ala
1745                1750                1755                1760

gtc gtc cga ctc gta cgc gaa cga ctc gcg cga ccg gcc tcc ccc gca   5328
Val Val Arg Leu Val Arg Glu Arg Leu Ala Arg Pro Ala Ser Pro Ala
               1765                1770                1775

cgc gat gtg gcc tcc gtc gtg gcc gag ttg gag tcc ctg ctg acg gcc   5376
Arg Asp Val Ala Ser Val Val Ala Glu Leu Glu Ser Leu Leu Thr Ala
           1780                1785                1790

ggt gcg gag gtc tcg gag gag acc gcg gcg cgg ttg aaa gcg gtg acg   5424
Gly Ala Glu Val Ser Glu Glu Thr Ala Ala Arg Leu Lys Ala Val Thr
       1795                1800                1805

gcg gtg tcg acg ggg acg acg ggg tcg ggc agc ggt acg ggc gcc ggc   5472
Ala Val Ser Thr Gly Thr Thr Gly Ser Gly Ser Gly Thr Gly Ala Gly
   1810                1815                1820

tcc ggc ggg gct ctg gac ctg gta tcg gcc agt gac gag gaa ctg ttc   5520
Ser Gly Gly Ala Leu Asp Leu Val Ser Ala Ser Asp Glu Glu Leu Phe
1825                1830                1835                1840

cgg ctg atg gac gcg gag agc tga                                  5544
Arg Leu Met Asp Ala Glu Ser
               1845
```

<210> 2
<211> 1848
<212> PRT
<213> Streptomyces natalensis

<400> 2

```
Met Val Pro Val His Thr Asp Asp Tyr Ala Ile Gln Pro Pro Ala Asp
  1               5                  10                  15

Thr Ala His Gly Gly Gly Gly Phe Thr Leu Pro Ala Val Phe Glu Ala
            20                  25                  30

Ala Val Glu Ser Ala Pro Asp Ala Val Ala Leu Val Asp Gly Thr Val
        35                  40                  45

Pro Gly Pro Gly Arg Met Trp Arg Ala Asp Val Asp Ala Leu Ala Arg
    50                  55                  60

Gly Leu Gln Glu Ser Gly Ile Ala Pro Gly Asp Val Val Ala Val Arg
65                  70                  75                  80

Leu Pro Asn Cys Gly Arg Phe Pro Thr Leu His Leu Ala Val Ala Ala
                85                  90                  95

Val Gly Ala Val Leu Leu Pro Ile His Gln Gly Thr Pro Leu Pro Glu
            100                 105                 110

Val Asp Ala Leu Leu Thr Arg Ala Glu Pro Ala Leu Leu Val Leu Ser
        115                 120                 125
```

Ala Ala Gly Ser Asp Gly Leu Ala Thr Ala Arg Ser Leu Leu Glu Ser
130 135 140

Val Pro Ser Leu Arg Gly Val Leu Leu Ala Gly Ala Ser Gly Asp Gly
145 150 155 160

Glu Ser Gly Ser Val Gly Gly Gly Glu Ser Gly Ser Gly Arg Arg Ser
165 170 175

Leu Asp Gly Leu Leu Ala Gly Trp Ala Gly Ser Gly Pro Arg Pro Val
180 185 190

Asp Val Thr Pro Asp Met Pro Leu Val Leu Val Pro Ser Ser Gly Thr
195 200 205

Val Ser Ala Arg Pro Lys Leu Cys Val His Ser His Asp Gly Leu Leu
210 215 220

Ser Asn Thr Ala Ala Val Thr Ala Glu Ala Ala Asp Ala Phe Asp Gly
225 230 235 240

Pro Val Leu Thr Ala Cys Pro Met Thr His Leu Phe Gly Leu Gln Ser
245 250 255

Leu His Ala Ala Leu Phe Ala Ala Cys Thr Gln Val Leu Leu Thr Gly
260 265 270

Trp Asp Val Asp Arg Phe Leu Glu Gln Ala Arg Glu His Gly Pro Arg
275 280 285

Val Val Phe Ala Val Pro Ala Gln Leu Arg Asp Val Val Thr Arg Leu
290 295 300

Ala Arg Thr Gly Glu Pro Ala Gly Phe Thr Pro Tyr Gln Val Arg Thr
305 310 315 320

Ala Gly Ala Ala Val Ala Pro Ala Leu Ala Val Arg Val Arg Ala Val
325 330 335

Leu Asp Cys Glu Leu Val Val Val Trp Gly Met Ser Glu Ile Gly Thr
340 345 350

Gly Thr Arg Thr Arg Ala His His Pro Asp Gly Cys Val Gly Glu Pro
355 360 365

Val Ser Gly Val Asp Val Arg Val Val Asp Glu His Gly Gln Glu Cys
370 375 380

Ala Ala Asp Glu Arg Gly Glu Leu Gln Tyr Arg Gly Pro Gly Leu Phe
385 390 395 400

Arg Gly Tyr Phe Arg Glu Pro Glu Leu Thr Arg Ser Ala Leu Thr Asp
405 410 415

Asp Gly Trp Leu Arg Thr Gly Asp Leu Ala Thr Val Asp Ala Asp Gly
420 425 430

Val Val Val Leu His Gly Arg Ala Ala Glu Leu Ile Asn Thr Gly Gly
435 440 445

Arg Lys Phe Ser Ala Gly Glu Val Glu Gly Leu Leu Ser Gly Phe Thr
450 455 460

Asp Leu Gly Pro Leu Ala Val Val Gly Ala Pro Asp Asp Arg Leu Gly

```
465                 470                 475                 480

Glu Tyr Pro Cys Leu Val Val Thr Asp His Ala Asp Gly Thr Ile Gly
                485                 490                 495

Leu Ser Glu Val Thr Ala Phe Leu Arg Arg Leu Gly Leu Ala Asp His
            500                 505                 510

Lys Ile Pro Leu Glu Leu Val Thr Val Arg Glu Leu Pro Phe Ser Pro
        515                 520                 525

Ala Gly Lys Leu Asp Arg Gly Ala Leu Lys Arg Leu Leu Ala Asn Leu
    530                 535                 540

Ala Glu Val Ser Val Pro Ala Arg Leu Gly Ala Val Pro Pro Tyr Thr
545                 550                 555                 560

Ala Glu Glu Ala Leu Asp Leu Val Arg Asp Cys Val Gly Arg Val Leu
                565                 570                 575

Arg Tyr Gly Gly Ala Ala Val Pro Phe Pro Pro Asp Lys Asp Phe Phe
            580                 585                 590

Ser Pro Asp Lys Asp Phe Arg Gln Leu Gly Leu Asp Ser Ile Gly Ala
        595                 600                 605

Val Arg Leu Arg Asn Leu Leu Arg Glu Glu Thr Gly Leu Pro Leu Pro
    610                 615                 620

Ala Thr Leu Ala Phe Asp Ser Pro Thr Pro Arg Ala Val Ala Arg Val
625                 630                 635                 640

Leu Ala Glu Gln Glu Glu Pro Ser Gln Asp Glu Pro Arg Glu Asn Pro
                645                 650                 655

Ala Asp Gly Ala Asp Pro Val Ala Ile Val Gly Met Ala Cys Arg Leu
            660                 665                 670

Pro Gly Gly Ala Asp Ser Pro Asp Ala Leu Trp Glu Leu Leu Ala Asp
        675                 680                 685

Gly Thr Asp Ala Met Ser Pro Phe Pro Thr Asp Arg Gly Trp Asp Leu
    690                 695                 700

Asp Arg Leu Phe Asp Glu Asp Ala Asp Arg Pro Gly Thr Ser Tyr Ala
705                 710                 715                 720

Arg Glu Gly Gly Phe Leu His Asp Ala Gly Asp Phe Asp Ala Gly Phe
                725                 730                 735

Phe Gly Leu Ser Asp Gln Glu Ala Thr Ala Thr Asp Pro Gln Gln Arg
            740                 745                 750

Leu Leu Leu Glu Ala Ala Trp Glu Thr Phe Glu Arg Ala Gly Ile Asp
        755                 760                 765

Pro Gln Ser Leu Arg Gly Ser Arg Thr Gly Val Phe Thr Gly Ala Met
    770                 775                 780

Asp Arg Gly Tyr Gly Thr Ser Ala Ser Ala Ala Pro Ser Ala Trp Glu
785                 790                 795                 800

Ser Met Leu Ile Thr Gly Thr Ala Gly Ser Ala Val Ser Gly Arg Ile
                805                 810                 815
```

Ala Tyr Thr Tyr Gly Leu Glu Gly Pro Ala Leu Thr Val Asp Thr Ala
820 825 830

Ser Ser Ser Ser Leu Val Ala Leu His Leu Ala Cys Arg Ser Leu Arg
835 840 845

Ser Gly Glu Thr Asp Leu Ala Leu Ala Gly Gly Val Thr Val Met Ala
850 855 860

Thr Pro Ala Pro Phe Ala His Phe Ser Arg Leu Arg Ala Leu Ser Pro
865 870 875 880

Asp Ser Arg Ser Met Ala Tyr Ala Asp Ala Ala Asn Gly Ser Ala Trp
885 890 895

Ser Glu Gly Ala Gly Leu Leu Leu Leu Glu Arg Leu Ser Asp Ala Arg
900 905 910

Arg Asn Gly His Arg Val Leu Ala Leu Val Arg Gly Ser Ala Val Asn
915 920 925

Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Ser Gly Pro Ala Gln
930 935 940

Gln Arg Val Ile Arg Gln Ala Leu Ala Asp Ala Gly Leu Thr Pro Gln
945 950 955 960

Asp Val Asp Ala Val Glu Gly His Gly Thr Gly Thr Pro Leu Gly Asp
965 970 975

Pro Ile Glu Ala Gln Ala Leu Leu Ala Thr Tyr Gly Gln Gln Arg Pro
980 985 990

Val Glu Arg Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Phe Gly His
995 1000 1005

Thr Gln Ala Ala Ala Gly Val Val Gly Val Ile Lys Thr Val Leu Ala
1010 1015 1020

Leu Arg His Gly Val Leu Pro Gln Thr Leu His Val Asp Ala Pro Ser
025 1030 1035 1040

Ala Lys Val Asp Trp Ser Ala Gly Ser Val Arg Leu Leu Thr Glu Ala
1045 1050 1055

Arg Pro Trp Pro Arg Glu Ser Gly Arg Thr Arg Arg Ala Gly Val Ser
1060 1065 1070

Ser Phe Gly Leu Thr Gly Thr Asn Ala His Val Ile Leu Glu Glu Ala
1075 1080 1085

Pro Gly Glu Ala Ala Ala Gly Ala Arg Ala Glu Val Pro Glu Glu Ala
1090 1095 1100

Arg Cys Ala Ser Ser Pro Ala Arg Leu Pro Glu Pro Pro Gly Asp Ala
105 1110 1115 1120

Ala Ala Pro Trp Val Leu Ser Ala Arg Ser Arg Ala Ala Leu Arg Ala
1125 1130 1135

Gln Ala Leu Arg Leu Ala Asp Gln Val Ala Ala Asp Pro Gly Leu Arg
1140 1145 1150

Ala Gln Asp Val Ala His Ala Leu Ala Thr Ser Arg Thr Leu His Arg
1155 1160 1165

His Arg Ala Val Val Ser Gly Ser Asp Arg Ala Gln Met Leu Ala Ala
1170 1175 1180

Ala Lys Arg Phe Gly Leu Gly Glu Arg Thr Ala Gly Val Thr Pro Asp
185 1190 1195 1200

Asp Ser Ala Pro Gly Leu Leu Ala Phe Val Phe Ser Gly Gln Gly Ser
1205 1210 1215

Gln Arg Ser Gly Met Gly Arg Ala Ala Ala Glu Ala Phe Pro Val Phe
1220 1225 1230

Gly Arg Ala Leu Gly Glu Val Cys Ala Ala Leu Asp Pro Leu Leu Thr
1235 1240 1245

Arg Pro Leu Thr Ser Val Met Trp Ala Ala Pro Gly Ser Glu Glu Ala
1250 1255 1260

Ala Arg Leu Asp Asp Thr Thr Tyr Thr Gln Pro Ala Leu Phe Ala Val
265 1270 1275 1280

Gln Val Ala Leu Tyr Arg Leu Phe Glu Ser Trp Gly Val Val Pro Asp
1285 1290 1295

Gln Leu Val Gly His Ser Val Gly Glu Ile Ser Ala Ala His Val Ala
1300 1305 1310

Gly Val Leu Gly Leu Arg Asp Ala Cys Thr Leu Val Ala Ala Arg Ser
1315 1320 1325

Arg Leu Met Gly Ala Leu Pro Pro Gly Gly Ala Met Val Ala Val Arg
1330 1335 1340

Ile Thr Glu Pro Glu Val Thr Pro Trp Leu Ala Glu Leu Thr Asp Glu
345 1350 1355 1360

Val Ser Ile Ala Ala Val Asn Gly Pro His Ser Leu Val Leu Ala Gly
1365 1370 1375

Ala Glu Ala Pro Leu Val Ala Leu Thr Asp Arg Leu Ala Ala Ala Gly
1380 1385 1390

His Lys Thr Arg Arg Leu Met Val Ser Thr Ala Pro His Ser Pro Leu
1395 1400 1405

Met Asp Pro Met Leu Glu Glu Phe Arg Ala Val Val Arg Thr Leu Ser
1410 1415 1420

Tyr Ala Ala Pro Ala Val Pro Leu Val Ser Thr Val Thr Gly Arg Pro
425 1430 1435 1440

Leu Thr Gly Glu Glu Ala Arg Asp Pro Asp His Trp Val Arg His Val
1445 1450 1455

Arg Gln Ser Val Arg Phe Lys Asp Ala Ile Gly Arg Leu Arg Asp Glu
1460 1465 1470

Arg Val Thr Gly Phe Leu Glu Leu Gly Ala Glu Pro Ala Leu Thr Pro
1475 1480 1485

Met Ile Asp Glu Cys Leu Glu Ser Ala Asp Gly Gln Pro Gly Thr Ala
1490 1495 1500

Leu Val Pro Ser Leu Arg Ala Gly Val Pro Glu Arg Asp Ala Leu Leu

505 1510 1515 1520

Thr Ala Val Ala Arg Val His Ala Gln Gly Val Pro Val Asp Trp Asp
1525 1530 1535

Ala Val Leu Pro Gly Ala Glu Ala Ser Val Thr Val Arg Gly Leu Pro
1540 1545 1550

Ala Ala Asp Arg Gln Trp Phe Arg Phe Val Pro Asp Gln Gly Ala Pro
1555 1560 1565

Leu Thr Leu Ala Asp Arg Ser Leu His Leu Glu Gly Ala Ala His Leu
1570 1575 1580

Arg Asp Val Gly Gly Cys Arg Thr Ala Asp Gly Arg Trp Val Lys Met
585 1590 1595 1600

Gly Val Leu Tyr Arg Thr Asn Asn Leu His Ala Leu Thr Asp Ala Asp
1605 1610 1615

Leu Ala Lys Leu Gln Arg Leu Gly Ile Arg Thr Asp Phe Asp Leu Arg
1620 1625 1630

Met Pro Gly Glu Arg Ala Lys Ala Pro Asn Arg Val Pro Thr Gly Ala
1635 1640 1645

Arg Tyr Ile Val Ala Asp Ala Phe Asp Ala His Leu Arg Asp Gly Leu
1650 1655 1660

Ser Leu Ser Glu Arg Asp Leu Glu Arg Leu Arg Glu Ala Arg Leu Ser
665 1670 1675 1680

Gly Leu Asp Glu Ala Glu Gln His Ala Leu Val Leu Ala Leu Val Leu
1685 1690 1695

Ala Glu Thr Ser Ala Val Leu Gly Gly Gln Glu Thr Pro Gly Glu Glu
1700 1705 1710

Pro His Gly Glu Glu Gly His Arg Thr Phe Lys Glu Met Gly Ile Asn
1715 1720 1725

Ser Leu Asn Ala Val Glu Leu Arg Asn Arg Leu Ile Ala Ala Thr Asp
1730 1735 1740

Leu Arg Leu Pro Ala Thr Leu Val Tyr Asp Tyr Pro Thr Pro Lys Ala
745 1750 1755 1760

Val Val Arg Leu Val Arg Glu Arg Leu Ala Arg Pro Ala Ser Pro Ala
1765 1770 1775

Arg Asp Val Ala Ser Val Val Ala Glu Leu Glu Ser Leu Leu Thr Ala
1780 1785 1790

Gly Ala Glu Val Ser Glu Glu Thr Ala Ala Arg Leu Lys Ala Val Thr
1795 1800 1805

Ala Val Ser Thr Gly Thr Thr Gly Ser Gly Ser Gly Thr Gly Ala Gly
1810 1815 1820

Ser Gly Gly Ala Leu Asp Leu Val Ser Ala Ser Asp Glu Glu Leu Phe
825 1830 1835 1840

Arg Leu Met Asp Ala Glu Ser
1845

<210> 3
<211> 20394
<212> DNA
<213> Streptomyces natalensis

<220>
<221> CDS
<222> (1)..(20394)

<400> 3

```
atg tcg aac gag gag aag ctg cgc gag tac ctc aag cgc gcg atc gcg   48
Met Ser Asn Glu Glu Lys Leu Arg Glu Tyr Leu Lys Arg Ala Ile Ala
  1               5                  10                  15

gac ctt cac gag act cgt cag caa ttg gac gag acc gag gcg aag cag   96
Asp Leu His Glu Thr Arg Gln Gln Leu Asp Glu Thr Glu Ala Lys Gln
             20                  25                  30

cga gag ccc ctc gcg atc gtg tcg atg gcc tgc cgc ttc ccc ggc ggc  144
Arg Glu Pro Leu Ala Ile Val Ser Met Ala Cys Arg Phe Pro Gly Gly
         35                  40                  45

gtc cgt tcg ccc gag gag ctg tgg gag ctg ctg cgc gac ggc gtc gac  192
Val Arg Ser Pro Glu Glu Leu Trp Glu Leu Leu Arg Asp Gly Val Asp
     50                  55                  60

gcg gtt tcc tcc ttc ccc cgt aac cgc ggc tgg gac ctg gac gcg ctc  240
Ala Val Ser Ser Phe Pro Arg Asn Arg Gly Trp Asp Leu Asp Ala Leu
 65                  70                  75                  80

tac cac tcc gac ccg gcc cac cag ggc acc agc tat gcg cgc gag ggc  288
Tyr His Ser Asp Pro Ala His Gln Gly Thr Ser Tyr Ala Arg Glu Gly
                 85                  90                  95

gga ttc ctg cat gac gcg ggc gag ttc gac ccc ggc ttc ttc ggg atc  336
Gly Phe Leu His Asp Ala Gly Glu Phe Asp Pro Gly Phe Phe Gly Ile
            100                 105                 110

tcc ccg cgc gag gcg ctc gcc atg gac ccc cag cag cgg ctg ctg ctg  384
Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg Leu Leu Leu
        115                 120                 125

gag acc gca tgg gaa gcc gtc gag cgg gcc ggt atc gac ccg gag tcc  432
Glu Thr Ala Trp Glu Ala Val Glu Arg Ala Gly Ile Asp Pro Glu Ser
    130                 135                 140

ctc gcg ggc agc cga acg ggt gtc ttt gtc ggc acc ggg cac gga ggg  480
Leu Ala Gly Ser Arg Thr Gly Val Phe Val Gly Thr Gly His Gly Gly
145                 150                 155                 160

tac gac gcc gag ggc cga cgg cgt gcc gac gag gtc ggc ggg cac ttg  528
Tyr Asp Ala Glu Gly Arg Arg Arg Ala Asp Glu Val Gly Gly His Leu
                165                 170                 175

ctg acg ggc aat cac atc agc atc gcc tcc ggc cgg att tcg tat gtc  576
Leu Thr Gly Asn His Ile Ser Ile Ala Ser Gly Arg Ile Ser Tyr Val
            180                 185                 190

ctg ggg ctg gaa ggc cct gcc ctg acc gtg gac acg gcc tgc tcc tcg  624
Leu Gly Leu Glu Gly Pro Ala Leu Thr Val Asp Thr Ala Cys Ser Ser
        195                 200                 205

tcg ctg gtc gcc ctg cat ctg gcg atg cac gcg ctg cgg cgc gac gaa  672
Ser Leu Val Ala Leu His Leu Ala Met His Ala Leu Arg Arg Asp Glu
```

```
     210                 215                 220

tgc gcc atg gcc ctg gtg ggc ggc gcg acc gtg atg tcc acg ccg cag   720
Cys Ala Met Ala Leu Val Gly Gly Ala Thr Val Met Ser Thr Pro Gln
225                 230                 235                 240

atg ttc gtg gag ttc tcc cgc cag cgc ggc ctt gcc ccc gac ggc cgc   768
Met Phe Val Glu Phe Ser Arg Gln Arg Gly Leu Ala Pro Asp Gly Arg
                245                 250                 255

tgc aag ccg ttc gcg gcc gcc gcc gac ggc acc ggc tgg agc gag ggt   816
Cys Lys Pro Phe Ala Ala Ala Ala Asp Gly Thr Gly Trp Ser Glu Gly
            260                 265                 270

gtc gga ctg ctg ctc gtc gag cgg ctc agt gac gcc gta cgc aac ggc   864
Val Gly Leu Leu Leu Val Glu Arg Leu Ser Asp Ala Val Arg Asn Gly
        275                 280                 285

tat ccc gtc ctc gcc gtg ctg aag ggc tcg gcc gtc aac cag gac ggc   912
Tyr Pro Val Leu Ala Val Leu Lys Gly Ser Ala Val Asn Gln Asp Gly
    290                 295                 300

gcg tcc aac ggc ctg acc gcc ccc aac ggc ccc tcg cag caa cgc gtc   960
Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ser Gln Gln Arg Val
305                 310                 315                 320

atc cgc cag gcg ctg acc ggc gcg ggc ctc gcc gcc tcg gac atc gac  1008
Ile Arg Gln Ala Leu Thr Gly Ala Gly Leu Ala Ala Ser Asp Ile Asp
                325                 330                 335

gcc gtg gag gcg cac ggc acc ggc acc acc ctc ggc gac ccc gtc gag  1056
Ala Val Glu Ala His Gly Thr Gly Thr Thr Leu Gly Asp Pro Val Glu
            340                 345                 350

gcg cac gcc ctg ctg gcc acc tac ggg cag cag cgc gcc gcc gac cgg  1104
Ala His Ala Leu Leu Ala Thr Tyr Gly Gln Gln Arg Ala Ala Asp Arg
        355                 360                 365

ccg tgc gga ctc ggc tcc atg aag tcc aac atc ggg cac acc cag gcc  1152
Pro Cys Gly Leu Gly Ser Met Lys Ser Asn Ile Gly His Thr Gln Ala
    370                 375                 380

gcc gcc ggt atc gcc ggc gtg atg aaa atg gtc ctg gcg atg cgg cac  1200
Ala Ala Gly Ile Ala Gly Val Met Lys Met Val Leu Ala Met Arg His
385                 390                 395                 400

ggg cac ctg ccc agg acc ctg cac ctg gac gag ccc acc ggg cac gtc  1248
Gly His Leu Pro Arg Thr Leu His Leu Asp Glu Pro Thr Gly His Val
                405                 410                 415

gac tgg agc gag ggc aac gcc agg ctc ctc gcg gag ccc gag ccc tgg  1296
Asp Trp Ser Glu Gly Asn Ala Arg Leu Leu Ala Glu Pro Glu Pro Trp
            420                 425                 430

ccg agc gcc ggc cgg ccc cgt cgc gcc gcc gtc tcc tcc ttc ggc atc  1344
Pro Ser Ala Gly Arg Pro Arg Arg Ala Ala Val Ser Ser Phe Gly Ile
        435                 440                 445

agc ggc acc aac gcc cac gtc atc ctg gag cag gcg ccc gcc cac gag  1392
Ser Gly Thr Asn Ala His Val Ile Leu Glu Gln Ala Pro Ala His Glu
    450                 455                 460

gcc gaa ccg gcc ccc gaa ccg gcc gcc cgg ccg ggc gcg ctg ccc tgg  1440
Ala Glu Pro Ala Pro Glu Pro Ala Ala Arg Pro Gly Ala Leu Pro Trp
465                 470                 475                 480
```

```
atc ctg tcc gcc cgc acc gaa gcg ggc ctg cgt gcc cag gcc gac cgc   1488
Ile Leu Ser Ala Arg Thr Glu Ala Gly Leu Arg Ala Gln Ala Asp Arg
                485                 490                 495

ctc ggc cgc cac cta cgg gac cgc gcc gac ctc gaa ccg gcc gcc gtc   1536
Leu Gly Arg His Leu Arg Asp Arg Ala Asp Leu Glu Pro Ala Ala Val
            500                 505                 510

gcc cat gcg ctc gcg gac acc cgg acc ctc atg gaa cac cgc gcg gta   1584
Ala His Ala Leu Ala Asp Thr Arg Thr Leu Met Glu His Arg Ala Val
        515                 520                 525

gtc gtc gcg ggc gac cgg gag gag ttc ctg cgc ggc ctg gac gcc ctc   1632
Val Val Ala Gly Asp Arg Glu Glu Phe Leu Arg Gly Leu Asp Ala Leu
    530                 535                 540

gcc gcg ggc cgc acc gcc aac ggc ctg gtc agc ggc gtc gcc gtc aag   1680
Ala Ala Gly Arg Thr Ala Asn Gly Leu Val Ser Gly Val Ala Val Lys
545                 550                 555                 560

gcc gcc agc gcg ttc ctc ttc gcc ggg cag ggc tcc cag cga ccg ggc   1728
Ala Ala Ser Ala Phe Leu Phe Ala Gly Gln Gly Ser Gln Arg Pro Gly
                565                 570                 575

atg ggg cgc gaa ctg cac gcc gcg cac ccc gtg ttc gcc acg gcc ttc   1776
Met Gly Arg Glu Leu His Ala Ala His Pro Val Phe Ala Thr Ala Phe
            580                 585                 590

gac gcg gtg tgc gcc gaa ctg gac cca cac ctg gac cgg ccg ctg cgc   1824
Asp Ala Val Cys Ala Glu Leu Asp Pro His Leu Asp Arg Pro Leu Arg
        595                 600                 605

gac atc gtc ttc gcc gag gaa ggc agc gcc gag gcc gcc ctg ctc gac   1872
Asp Ile Val Phe Ala Glu Glu Gly Ser Ala Glu Ala Ala Leu Leu Asp
    610                 615                 620

cag acc gcc tac acc cag gcc gcg ctc ttc gcc ctg gaa acc gcc ctg   1920
Gln Thr Ala Tyr Thr Gln Ala Ala Leu Phe Ala Leu Glu Thr Ala Leu
625                 630                 635                 640

ttc cgg ctc gtc gaa tcc tgg ggc gtg gca ccc cgg ttc gtc gcc gga   1968
Phe Arg Leu Val Glu Ser Trp Gly Val Ala Pro Arg Phe Val Ala Gly
                645                 650                 655

cac tcc atc ggc gag ctg acc gcc gcc cac gtc agt ggc gtg ctg acc   2016
His Ser Ile Gly Glu Leu Thr Ala Ala His Val Ser Gly Val Leu Thr
            660                 665                 670

ctc cac gac gcc gca cgg ctg gtc gcc gcg cgc ggc acc ctc atg cag   2064
Leu His Asp Ala Ala Arg Leu Val Ala Ala Arg Gly Thr Leu Met Gln
        675                 680                 685

gcg ctg ccc gca ggc ggc gcc atg gtg gcg gtc cag gcc acc gag gac   2112
Ala Leu Pro Ala Gly Gly Ala Met Val Ala Val Gln Ala Thr Glu Asp
    690                 695                 700

gag atc cgc gag cgt ctc gcc ggc cac gag gac cac gtc gcc ctc gcg   2160
Glu Ile Arg Glu Arg Leu Ala Gly His Glu Asp His Val Ala Leu Ala
705                 710                 715                 720

gcc gcc aac ggg ccc gat tcc acc gtc att tcg ggc gac gaa cag gcc   2208
Ala Ala Asn Gly Pro Asp Ser Thr Val Ile Ser Gly Asp Glu Gln Ala
                725                 730                 735
```

```
gtc acc gag atc gcg gcc cac tgg gag gca cag ggc cgc cgc acc aag   2256
Val Thr Glu Ile Ala Ala His Trp Glu Ala Gln Gly Arg Arg Thr Lys
            740                 745                 750

cgg ctg cgg gtc agc cac gcc ttc cac tcc ccg cac atg gac gac atg   2304
Arg Leu Arg Val Ser His Ala Phe His Ser Pro His Met Asp Asp Met
        755                 760                 765

ctg gag gac ttc cgg cgc gtc gcc cgc ggt ctg acc ttc cac gcc ccc   2352
Leu Glu Asp Phe Arg Arg Val Ala Arg Gly Leu Thr Phe His Ala Pro
    770                 775                 780

cgc atc ccc gtg gtg tcc acg gtg acc ggc gcg ctc gcc acc gaa gac   2400
Arg Ile Pro Val Val Ser Thr Val Thr Gly Ala Leu Ala Thr Glu Asp
785                 790                 795                 800

gaa ctg cgc tcg ccc gac tac tgg gtg cgg cag gtc cgc gaa acc gtc   2448
Glu Leu Arg Ser Pro Asp Tyr Trp Val Arg Gln Val Arg Glu Thr Val
                805                 810                 815

cgc ttc tgt gct gcc gtg cgc acc ctt gag gcc gag ggc gtc acc acc   2496
Arg Phe Cys Ala Ala Val Arg Thr Leu Glu Ala Glu Gly Val Thr Thr
            820                 825                 830

ttc gtg gag atc ggc acc ggc ggc gtc ctc acc ccc atg gtc cag gac   2544
Phe Val Glu Ile Gly Thr Gly Gly Val Leu Thr Pro Met Val Gln Asp
        835                 840                 845

tgt ctg acc acc ctc gaa gag ccc gtt ctc gtc cct ctg ctg cgc acc   2592
Cys Leu Thr Thr Leu Glu Glu Pro Val Leu Val Pro Leu Leu Arg Thr
    850                 855                 860

ggc cgc ccc gaa acc gtc gct ctc acc gag ggc gtc gcc acc gcc ttc   2640
Gly Arg Pro Glu Thr Val Ala Leu Thr Glu Gly Val Ala Thr Ala Phe
865                 870                 875                 880

gtg cac ggt gtc ccc gtc gac cgg tcc gcc ttc ccg ggc gcg ccc ggt   2688
Val His Gly Val Pro Val Asp Arg Ser Ala Phe Pro Gly Ala Pro Gly
                885                 890                 895

acc tcc cgc gcg gac ctg ccc acc tac gcc ttc cag cgt cag tgg tac   2736
Thr Ser Arg Ala Asp Leu Pro Thr Tyr Ala Phe Gln Arg Gln Trp Tyr
            900                 905                 910

tgg ctg gac ccg gcc gac cac gac gag ggg gag gcg gcc gcc gcc gaa   2784
Trp Leu Asp Pro Ala Asp His Asp Glu Gly Glu Ala Ala Ala Ala Glu
        915                 920                 925

gcg ggc gag gcc gga ttc tgg gcg gcc gtc gaa cgc gag gac ctc cag   2832
Ala Gly Glu Ala Gly Phe Trp Ala Ala Val Glu Arg Glu Asp Leu Gln
    930                 935                 940

gag ctg tcg gcc gtc ctg gcc atc gac ggc agc gaa gcg gac tcc ctc   2880
Glu Leu Ser Ala Val Leu Ala Ile Asp Gly Ser Glu Ala Asp Ser Leu
945                 950                 955                 960

ggc agc ttc ctg ccc acc ctc tcc tcc tgg cgc agg cag cgc agg acc   2928
Gly Ser Phe Leu Pro Thr Leu Ser Ser Trp Arg Arg Gln Arg Arg Thr
                965                 970                 975

cag gcc gcc gcg gac cgc ttc agc tac cgc acc cac tgg gcc ccg cgc   2976
Gln Ala Ala Ala Asp Arg Phe Ser Tyr Arg Thr His Trp Ala Pro Arg
            980                 985                 990

acc gcc tcg ggc ggc ccc acc gcc acc ggg cac tgg ctc gtc gtc ctg   3024
```

```
Thr Ala Ser Gly Gly Pro Thr Ala Thr Gly His Trp Leu Val Val Leu
        995                 1000                1005

ccc gaa ggc ggc acc gac gac ccg tgg acc gcc cgc ctc ctg gac gcg   3072
Pro Glu Gly Gly Thr Asp Asp Pro Trp Thr Ala Arg Leu Leu Asp Ala
   1010                1015                1020

ctg aac gac cag ggc ctg cac acc gac gta cgc gaa ctg ccc gcc gac   3120
Leu Asn Asp Gln Gly Leu His Thr Asp Val Arg Glu Leu Pro Ala Asp
1025                1030                1035                1040

cac gag ccc gac gcc tgg ggc cga cac ccc gtg gac ggc gtg ctc tgt   3168
His Glu Pro Asp Ala Trp Gly Arg His Pro Val Asp Gly Val Leu Cys
                1045                1050                1055

ctg ctg gca ctc gac gag cgg ccc acc cgc tcc tgc cct ccg tac cgg   3216
Leu Leu Ala Leu Asp Glu Arg Pro Thr Arg Ser Cys Pro Pro Tyr Arg
            1060                1065                1070

cgc ggg ctg gcc gcc acc acc aac gct gct gcg cgc cct gag ggc gcg   3264
Arg Gly Leu Ala Ala Thr Thr Asn Ala Ala Ala Arg Pro Glu Gly Ala
        1075                1080                1085

ggc atc cag gca ccg ctg tgg tgc gtg acc cgc ggc gcc gtc gcc gtc   3312
Gly Ile Gln Ala Pro Leu Trp Cys Val Thr Arg Gly Ala Val Ala Val
    1090                1095                1100

gac cgg cac gag gcg ctc aag agc ccc tta cag gca cag aca tgg ggc   3360
Asp Arg His Glu Ala Leu Lys Ser Pro Leu Gln Ala Gln Thr Trp Gly
1105                1110                1115                1120

ctg ggc cgg gtg gcc gcc ctg gag tcc ccg cag agc tgg ggc ggg ctc   3408
Leu Gly Arg Val Ala Ala Leu Glu Ser Pro Gln Ser Trp Gly Gly Leu
                1125                1130                1135

atc gac ctg ccc gac aac ctg gac gga cgg gcc gtc tcc gcg ctg ctg   3456
Ile Asp Leu Pro Asp Asn Leu Asp Gly Arg Ala Val Ser Ala Leu Leu
            1140                1145                1150

agc acc ctc gcc ggg gag gag gac cag gtc gcc gtc cgc ccc gcc ggg   3504
Ser Thr Leu Ala Gly Glu Glu Asp Gln Val Ala Val Arg Pro Ala Gly
        1155                1160                1165

gtc ttc gcc cgc cgc ctg gaa cgg atc aca ccc ggc ggc gac acc ggc   3552
Val Phe Ala Arg Arg Leu Glu Arg Ile Thr Pro Gly Gly Asp Thr Gly
    1170                1175                1180

gac cgg tgg agc acc cac ggc acc gtc ctg gtc acc ggc ggc acc ggt   3600
Asp Arg Trp Ser Thr His Gly Thr Val Leu Val Thr Gly Gly Thr Gly
1185                1190                1195                1200

gcc ctc ggc gcg cac ctc gcc cac tgg ctg gcc gac gcc gga gcc gaa   3648
Ala Leu Gly Ala His Leu Ala His Trp Leu Ala Asp Ala Gly Ala Glu
                1205                1210                1215

cac ctc gtg ctc acc ggc cgc cgc ggc ccg cag gcc ccc ggc gca ccg   3696
His Leu Val Leu Thr Gly Arg Arg Gly Pro Gln Ala Pro Gly Ala Pro
            1220                1225                1230

gaa ctc gcg gcc gcc ctc acc gac cgg ggc gtc aag gtc acc ctc gcc   3744
Glu Leu Ala Ala Ala Leu Thr Asp Arg Gly Val Lys Val Thr Leu Ala
        1235                1240                1245

gcc tgc gac gcc gcc gac cgt gat gcg ctg gcg gcc gtc ctc gcg gac   3792
Ala Cys Asp Ala Ala Asp Arg Asp Ala Leu Ala Ala Val Leu Ala Asp
```

```
   1250                1255                1260

atc ccg ccg cac ctg ccg ctg acc ggc gtc gtc cac gcc gcg ggc gta   3840
Ile Pro Pro His Leu Pro Leu Thr Gly Val Val His Ala Ala Gly Val
1265               1270                1275                1280

ctg gac gac ggc gta ctg gac gcg ctc acc ccc gag cgc ttc gag acc   3888
Leu Asp Asp Gly Val Leu Asp Ala Leu Thr Pro Glu Arg Phe Glu Thr
                1285                1290                1295

gta ctg cgc ccc aag gcg cgg gcc gca cag aac ctg cac gaa ctc acc   3936
Val Leu Arg Pro Lys Ala Arg Ala Ala Gln Asn Leu His Glu Leu Thr
            1300                1305                1310

cag gac ctc gac ctg gac cac ttc gtg ctg ttc tcc tcg atc gtc ggc   3984
Gln Asp Leu Asp Leu Asp His Phe Val Leu Phe Ser Ser Ile Val Gly
        1315                1320                1325

gtc ctg ggc aac gcc gga cag gcc aac tac gcc gcc gcc aac gcc tac   4032
Val Leu Gly Asn Ala Gly Gln Ala Asn Tyr Ala Ala Ala Asn Ala Tyr
    1330                1335                1340

ttg gac gcc ctc gcc gaa cac cgt ctc gcc cag ggg ctc ccg gcc acc   4080
Leu Asp Ala Leu Ala Glu His Arg Leu Ala Gln Gly Leu Pro Ala Thr
1345                1350                1355                1360

tcc gtg tcc tgg ggc cct ggg cag gcg gcg gca tgg cac gac agc gac   4128
Ser Val Ser Trp Gly Pro Gly Gln Ala Ala Ala Trp His Asp Ser Asp
                1365                1370                1375

gcc gcc gac cgg atg agc cgc gac gga ctg ctg ccc atg gcc gcg gcc   4176
Ala Ala Asp Arg Met Ser Arg Asp Gly Leu Leu Pro Met Ala Ala Ala
            1380                1385                1390

ccg cgt cgc cgc cct gcg cca gcc ctc gcc cag ggc atg aca cag gtg   4224
Pro Arg Arg Arg Pro Ala Pro Ala Leu Ala Gln Gly Met Thr Gln Val
        1395                1400                1405

acc gtg gcc gac atc gac tgg agc gca tac gcc ccc gcc ctg acc gcc   4272
Thr Val Ala Asp Ile Asp Trp Ser Ala Tyr Ala Pro Ala Leu Thr Ala
    1410                1415                1420

gtc cgc ccc agc ccc ctc atc ggc gac ctg ccc gag gca cgc cgc gcg   4320
Val Arg Pro Ser Pro Leu Ile Gly Asp Leu Pro Glu Ala Arg Arg Ala
1425                1430                1435                1440

ctc ggc ccc gca gaa ggc ccc cgc cgg gaa cgc tcc ccc ctg cgc gac   4368
Leu Gly Pro Ala Glu Gly Pro Arg Arg Glu Arg Ser Pro Leu Arg Asp
                1445                1450                1455

cgg atc ggc gca ctg ccg ccc gcc gaa cag gaa aag gca ttc ctg acc   4416
Arg Ile Gly Ala Leu Pro Pro Ala Glu Gln Glu Lys Ala Phe Leu Thr
            1460                1465                1470

atg gtc agg gaa gag gcc gcg agg gta ctg gga cac ccc tcg ccg gac   4464
Met Val Arg Glu Glu Ala Ala Arg Val Leu Gly His Pro Ser Pro Asp
        1475                1480                1485

acc gtc gat gcc caa cgc gcc ttc cgc gag cag ggg ttc gac tcc ctg   4512
Thr Val Asp Ala Gln Arg Ala Phe Arg Glu Gln Gly Phe Asp Ser Leu
    1490                1495                1500

atg gcc gtc gac ctg cgc aac cgg ctc tcc gcc gcg acg ggc ctg cgg   4560
Met Ala Val Asp Leu Arg Asn Arg Leu Ser Ala Ala Thr Gly Leu Arg
1505                1510                1515                1520
```

```
ctg ccc gcc acc ctg ctg ttc gac cac ccc acc ccc ctt gcg gcc gcc   4608
Leu Pro Ala Thr Leu Leu Phe Asp His Pro Thr Pro Leu Ala Ala Ala
                1525                1530                1535

gcc tgc ctg cgc tcc gaa gtc ctg ggc gcc gca gga ccc gcc acg gtc   4656
Ala Cys Leu Arg Ser Glu Val Leu Gly Ala Ala Gly Pro Ala Thr Val
            1540                1545                1550

gtt cag gca tcg acc gcc gcc ctc gac gaa ccg gtg gcg atc atc ggc   4704
Val Gln Ala Ser Thr Ala Ala Leu Asp Glu Pro Val Ala Ile Ile Gly
        1555                1560                1565

atg gcc tgc cgc ttc ccc ggc ggc gtg cac tca ccc gag gcc ctg tgg   4752
Met Ala Cys Arg Phe Pro Gly Gly Val His Ser Pro Glu Ala Leu Trp
    1570                1575                1580

cgg ctg ctg gcc gag ggc ggc gac gcc atc acc ccc atg ccc gcc gac   4800
Arg Leu Leu Ala Glu Gly Gly Asp Ala Ile Thr Pro Met Pro Ala Asp
1585                1590                1595                1600

cgg ggc tgg gac ctg gac cgg ctc tac cac ccc gac ccc gac cac cag   4848
Arg Gly Trp Asp Leu Asp Arg Leu Tyr His Pro Asp Pro Asp His Gln
                1605                1610                1615

ggc acc agc tac gcc cgc ggc ggc ggc ttc ctg gac ggc gcg gcc gac   4896
Gly Thr Ser Tyr Ala Arg Gly Gly Gly Phe Leu Asp Gly Ala Ala Asp
            1620                1625                1630

ttc gac gcg gac ttc ttc ggc atc tcg ccg cgc gag gcc ctc gcc atg   4944
Phe Asp Ala Asp Phe Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met
        1635                1640                1645

gac ccg cag cag cgg ctg ctc ctg gaa aca tgg gag gtg ctc gaa cag   4992
Asp Pro Gln Gln Arg Leu Leu Leu Glu Thr Trp Glu Val Leu Glu Gln
    1650                1655                1660

gcg ggg atc gac ccg gaa tcc ctg cgg ggc agc agc acc ggt gtc ttc   5040
Ala Gly Ile Asp Pro Glu Ser Leu Arg Gly Ser Ser Thr Gly Val Phe
1665                1670                1675                1680

gcg ggc acc aac acc cag gac tac ggc acg gcc ctg gac gcg gca cag   5088
Ala Gly Thr Asn Thr Gln Asp Tyr Gly Thr Ala Leu Asp Ala Ala Gln
                1685                1690                1695

gac gaa gcc ggc gga cac cgg ctc acc ggc aac gcg atg agc gtc gtc   5136
Asp Glu Ala Gly Gly His Arg Leu Thr Gly Asn Ala Met Ser Val Val
            1700                1705                1710

tcc ggc cgg gtc tcc tac acc ttc ggc ttc gag gga ccg gcc ctc acc   5184
Ser Gly Arg Val Ser Tyr Thr Phe Gly Phe Glu Gly Pro Ala Leu Thr
        1715                1720                1725

gtg gac acg gcg tgc tcc tcc tcg ctg gtg gcc ctg cac atg gcc gcg   5232
Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala Leu His Met Ala Ala
    1730                1735                1740

cag gcg ctg cgc cag ggc gaa tgc tcc ctg gcg gtc gcg ggc ggt gtg   5280
Gln Ala Leu Arg Gln Gly Glu Cys Ser Leu Ala Val Ala Gly Gly Val
1745                1750                1755                1760

acg gtg atg gcc acc ccg tcc tcc ttc gtg gag ttc gcc cgg cag cgc   5328
Thr Val Met Ala Thr Pro Ser Ser Phe Val Glu Phe Ala Arg Gln Arg
                1765                1770                1775
```

```
ggg ctg gcc ccc gac ggc cgc tgc aag ccg ttc gcg gcg gcc gcc gac   5376
Gly Leu Ala Pro Asp Gly Arg Cys Lys Pro Phe Ala Ala Ala Ala Asp
            1780                1785                1790

ggc acc ggc tgg agc gag ggc gtc ggc ctg ctg ctc gtg gaa cgg ctc   5424
Gly Thr Gly Trp Ser Glu Gly Val Gly Leu Leu Leu Val Glu Arg Leu
        1795                1800                1805

agc gac gcc cgc cga aac ggc cac cag gtg ctc gcc gtc gtc cgc ggt   5472
Ser Asp Ala Arg Arg Asn Gly His Gln Val Leu Ala Val Val Arg Gly
    1810                1815                1820

tcg gcg gtc aac cag gac ggc gcg tcc aac ggt ctg agc gca ccc agc   5520
Ser Ala Val Asn Gln Asp Gly Ala Ser Asn Gly Leu Ser Ala Pro Ser
1825                1830                1835                1840

ggc ccg tcc cag cag cgg gtg atc cgg cag gcc ctg gcg aac gcc cgg   5568
Gly Pro Ser Gln Gln Arg Val Ile Arg Gln Ala Leu Ala Asn Ala Arg
                1845                1850                1855

gtg gcc gcc tcc gag gtc gac gcc gtg gag gcc cac ggc acg ggc acc   5616
Val Ala Ala Ser Glu Val Asp Ala Val Glu Ala His Gly Thr Gly Thr
            1860                1865                1870

acg ctc ggt gac ccg atc gag gcc cag gcg ctg ctg gcc acc tac ggc   5664
Thr Leu Gly Asp Pro Ile Glu Ala Gln Ala Leu Leu Ala Thr Tyr Gly
        1875                1880                1885

cag gag cgg ccg ctg ctg ctc ggc gcg gtg aag tcc aac ctc ggc cac   5712
Gln Glu Arg Pro Leu Leu Leu Gly Ala Val Lys Ser Asn Leu Gly His
    1890                1895                1900

acc cag gcc gcc gcc ggt gtg gcg ggc gtg atg aag atg gtg ctg gcg   5760
Thr Gln Ala Ala Ala Gly Val Ala Gly Val Met Lys Met Val Leu Ala
1905                1910                1915                1920

atg cgg cac ggc atg ctg ccg cgc acc ctg cac gtc gac gag ccc acc   5808
Met Arg His Gly Met Leu Pro Arg Thr Leu His Val Asp Glu Pro Thr
                1925                1930                1935

ggg cat gtc gac tgg acc gcg ggc gcg gtc gag ctg ctc acc gag cac   5856
Gly His Val Asp Trp Thr Ala Gly Ala Val Glu Leu Leu Thr Glu His
            1940                1945                1950

acg gac tgg ccc gag acc ggc cac ccc cgg cgc gcc gcg gtc tcc gcg   5904
Thr Asp Trp Pro Glu Thr Gly His Pro Arg Arg Ala Ala Val Ser Ala
        1955                1960                1965

ttc ggc atc agc ggc acc aat gcg cac gtg gtg ctg gaa ctg ccc gca   5952
Phe Gly Ile Ser Gly Thr Asn Ala His Val Val Leu Glu Leu Pro Ala
    1970                1975                1980

gcc gaa cag ccc ttg gtc gaa cag ccc tcg gcc gcg gag ccc gac gcg   6000
Ala Glu Gln Pro Leu Val Glu Gln Pro Ser Ala Ala Glu Pro Asp Ala
1985                1990                1995                2000

ccg gcc acc gct ccc gac cgg acg ccc acc gcc tcc gac ggg acg gcg   6048
Pro Ala Thr Ala Pro Asp Arg Thr Pro Thr Ala Ser Asp Gly Thr Ala
                2005                2010                2015

ccg ctg ctg ctc tcc gcc aag agc gag agc gcc ctg cgc gcc cag gcg   6096
Pro Leu Leu Leu Ser Ala Lys Ser Glu Ser Ala Leu Arg Ala Gln Ala
            2020                2025                2030

gcc cgg ctg cac tcc cac ctg gag cgc gac ccc gcg ctc cgg ctc acg   6144
```

```
Ala Arg Leu His Ser His Leu Glu Arg Asp Pro Ala Leu Arg Leu Thr
       2035                 2040                 2045

gac gcc gcg tac acg ctg atg acg cac cgc acg gcc ttc gcc cac cgc   6192
Asp Ala Ala Tyr Thr Leu Met Thr His Arg Thr Ala Phe Ala His Arg
   2050                 2055                 2060

gcg gcc gtc cgc gcc gcc gac cac gaa gcc gcg ctg cgc gcc ctg acc   6240
Ala Ala Val Arg Ala Ala Asp His Glu Ala Ala Leu Arg Ala Leu Thr
2065                2070                 2075                 2080

gcc ctg gct gcg ggc gag gcc gac ccc gcc gtg gac acc ggc acc gcc   6288
Ala Leu Ala Ala Gly Glu Ala Asp Pro Ala Val Asp Thr Gly Thr Ala
                2085                 2090                 2095

cac acc ggc cgg gac gcc gtc ctc ttc tcc ggc cag gga tcg caa cgc   6336
His Thr Gly Arg Asp Ala Val Leu Phe Ser Gly Gln Gly Ser Gln Arg
            2100                 2105                 2110

atc gga atg ggc cgg gag ttg tcc ggc cgc tac ccg gtg ttc gca gag   6384
Ile Gly Met Gly Arg Glu Leu Ser Gly Arg Tyr Pro Val Phe Ala Glu
       2115                 2120                 2125

gcc ttc gac acc gtg tgc gcg gcc ttg gac gag cat ctg gac cgc ccc   6432
Ala Phe Asp Thr Val Cys Ala Ala Leu Asp Glu His Leu Asp Arg Pro
   2130                 2135                 2140

ctg cgg gac gtg gtc cgg ggc gag gac gag gag ctg ctg aac cgg acc   6480
Leu Arg Asp Val Val Arg Gly Glu Asp Glu Glu Leu Leu Asn Arg Thr
2145                2150                 2155                 2160

gtc tac gcc cag gcg ggg ctg ttc gcc atc gag gtg gcc ctc ttc cgg   6528
Val Tyr Ala Gln Ala Gly Leu Phe Ala Ile Glu Val Ala Leu Phe Arg
                2165                 2170                 2175

ctc gtg gag tcc tgg ggc gta cgg ccg cac tac gtg gcc ggg cat tcc   6576
Leu Val Glu Ser Trp Gly Val Arg Pro His Tyr Val Ala Gly His Ser
            2180                 2185                 2190

gtc ggc gag atc gcc gcc gcg cac gtc gcc ggg gtg ttc tcg ctg gcc   6624
Val Gly Glu Ile Ala Ala Ala His Val Ala Gly Val Phe Ser Leu Ala
       2195                 2200                 2205

gat gcc tgc gcg ctg gtg gcg gca cgc gga cgg ctg atg cag gcg ctg   6672
Asp Ala Cys Ala Leu Val Ala Ala Arg Gly Arg Leu Met Gln Ala Leu
   2210                 2215                 2220

ccc gcc ggc ggc gcg atg gcg gcg atc cgg gcg acg gag gac gaa gtc   6720
Pro Ala Gly Gly Ala Met Ala Ala Ile Arg Ala Thr Glu Asp Glu Val
2225                2230                 2235                 2240

ctc ccg cac ctg gcg gac agc gtc tcg atc gcg gcc gtc aac ggc ccg   6768
Leu Pro His Leu Ala Asp Ser Val Ser Ile Ala Ala Val Asn Gly Pro
                2245                 2250                 2255

tcg tcg gtc gtc gtc tcc ggc gcc gag cac gcc gtg ctc tcc atc gcc   6816
Ser Ser Val Val Val Ser Gly Ala Glu His Ala Val Leu Ser Ile Ala
            2260                 2265                 2270

gcg cac ttc gag ggc gcg ggc cgc aag acc acc agg ctg cgg gtc tcg   6864
Ala His Phe Glu Gly Ala Gly Arg Lys Thr Thr Arg Leu Arg Val Ser
       2275                 2280                 2285

cac gcc ttc cac tcc ccg ctc atg gac ccg atg ctg gcc gac ttc cgc   6912
His Ala Phe His Ser Pro Leu Met Asp Pro Met Leu Ala Asp Phe Arg
```

```
   2290                2295                2300

gcc gtc gcc gag ggc ctg acc tac ggc gag ccg gag ctg gcc gtc gta   6960
Ala Val Ala Glu Gly Leu Thr Tyr Gly Glu Pro Glu Leu Ala Val Val
2305            2310                2315                2320

tcg aac gtc acc ggc caa ctc gcc acc ccg gac cag ctg cgc acc ccc   7008
Ser Asn Val Thr Gly Gln Leu Ala Thr Pro Asp Gln Leu Arg Thr Pro
                2325                2330                2335

gag tac tgg gtg acc cat gtc cgc gcg gcg gtg cgc ttc gcg gac ggg   7056
Glu Tyr Trp Val Thr His Val Arg Ala Ala Val Arg Phe Ala Asp Gly
            2340                2345                2350

ata cgg gct ctg ggg gcg gaa ggg gtg acg cgg ttc ctc gaa ctc ggc   7104
Ile Arg Ala Leu Gly Ala Glu Gly Val Thr Arg Phe Leu Glu Leu Gly
        2355                2360                2365

ccg gac ggc gtc ctg tcg gcc ttg gcc agg gag tcg gca ccg gac gac   7152
Pro Asp Gly Val Leu Ser Ala Leu Ala Arg Glu Ser Ala Pro Asp Asp
    2370                2375                2380

gcc gtg tgc act ccc gtg ctg cgc aag gac cgc tcc gag gcg gcg acc   7200
Ala Val Cys Thr Pro Val Leu Arg Lys Asp Arg Ser Glu Ala Ala Thr
2385                2390                2395                2400

ctc ctc gcg gcc ctg acg cac ctg cac gta cac gga acc gag atc gac   7248
Leu Leu Ala Ala Leu Thr His Leu His Val His Gly Thr Glu Ile Asp
                2405                2410                2415

tgg acc gcg ttc ctc gcc ggc cgc gac gcg cac gcc gtc gac ctg ccc   7296
Trp Thr Ala Phe Leu Ala Gly Arg Asp Ala His Ala Val Asp Leu Pro
            2420                2425                2430

acg tac gcc ttc cag cac cag cgg ttc tgg ccg acc ccc gac cac acc   7344
Thr Tyr Ala Phe Gln His Gln Arg Phe Trp Pro Thr Pro Asp His Thr
        2435                2440                2445

cgc acc ggt gac ctg ggc gcc gtc ggc ctc gaa gcg acc ggg cac ccg   7392
Arg Thr Gly Asp Leu Gly Ala Val Gly Leu Glu Ala Thr Gly His Pro
    2450                2455                2460

ctg ctg agc gcc gcc gtg gaa ctg ccg gac ggt gag ggc ctg ttg ttc   7440
Leu Leu Ser Ala Ala Val Glu Leu Pro Asp Gly Glu Gly Leu Leu Phe
2465                2470                2475                2480

acc acc cgc ctc tcg ctc cag acc cac ccc tgg ctg gcc ggg cac gtc   7488
Thr Thr Arg Leu Ser Leu Gln Thr His Pro Trp Leu Ala Gly His Val
                2485                2490                2495

gtc atg ggc tcg gtc ctg ctg ccg ggg acg gcc ttc gcc gaa ctc gcc   7536
Val Met Gly Ser Val Leu Leu Pro Gly Thr Ala Phe Ala Glu Leu Ala
            2500                2505                2510

ctc cgc gcc gcc gac gag gtg ggc tgc gac cgc gtc gac gaa ctg acc   7584
Leu Arg Ala Ala Asp Glu Val Gly Cys Asp Arg Val Asp Glu Leu Thr
        2515                2520                2525

ctg gcc gcc ccg ctc gtc ctg ccc gag cac ggc ggc gta cag ctc cag   7632
Leu Ala Ala Pro Leu Val Leu Pro Glu His Gly Gly Val Gln Leu Gln
    2530                2535                2540

ctg cgg gtg ggc ccc gcc gac gcg tcc ggc cgc cgc acc ctg acc gcc   7680
Leu Arg Val Gly Pro Ala Asp Ala Ser Gly Arg Arg Thr Leu Thr Ala
2545                2550                2555                2560
```

```
cgc tcc agg gcg gag ggc gac ggc gac cgc ccg tgg gtc cag cac gcc   7728
Arg Ser Arg Ala Glu Gly Asp Gly Asp Arg Pro Trp Val Gln His Ala
                2565                2570                2575

acc ggc gtc ctc gcg gaa ggg gag tcg acg ccc gaa ccc ggc tac gac   7776
Thr Gly Val Leu Ala Glu Gly Glu Ser Thr Pro Glu Pro Gly Tyr Asp
            2580                2585                2590

ttc cac acc gag tcc tgg ccg ccc gcc gac gcc gcg ccc gtc gaa ctg   7824
Phe His Thr Glu Ser Trp Pro Pro Ala Asp Ala Ala Pro Val Glu Leu
        2595                2600                2605

tcc ggc ctc tac ccg gac ttc gcc gca cac ggt ttc gac tac ggt ccc   7872
Ser Gly Leu Tyr Pro Asp Phe Ala Ala His Gly Phe Asp Tyr Gly Pro
    2610                2615                2620

cac ttc cag ggg ctg cgg acc gcc tgg cgc cga ggc gac gag gtg ttc   7920
His Phe Gln Gly Leu Arg Thr Ala Trp Arg Arg Gly Asp Glu Val Phe
2625                2630                2635                2640

gcc gag gtc gcc ctg ccc gcc gag gcc gaa ggc gag gca tcc gcg tac   7968
Ala Glu Val Ala Leu Pro Ala Glu Ala Glu Gly Glu Ala Ser Ala Tyr
                2645                2650                2655

gga ctc cat ccg gcg ctg ctc gac gcc gcc ctg cac gtc gtc gcg ttc   8016
Gly Leu His Pro Ala Leu Leu Asp Ala Ala Leu His Val Val Ala Phe
            2660                2665                2670

aac gga gtg gac cgc ggc gtc gtg ccg ttc tcc tgg gag agc gtc gcg   8064
Asn Gly Val Asp Arg Gly Val Val Pro Phe Ser Trp Glu Ser Val Ala
        2675                2680                2685

ctg cac gcc acc ggc gcc tcg gcc gta cgg atc cgg gtc gtc cgg cac   8112
Leu His Ala Thr Gly Ala Ser Ala Val Arg Ile Arg Val Val Arg His
    2690                2695                2700

agc ggc gac acg gtc tcc gtg gat gtc gcc gac acc acc ggc gag ccc   8160
Ser Gly Asp Thr Val Ser Val Asp Val Ala Asp Thr Thr Gly Glu Pro
2705                2710                2715                2720

gtc gcc tcc atc ggc acg ctc gtc ctg cgg gcg gtc tcc gcc gac cag   8208
Val Ala Ser Ile Gly Thr Leu Val Leu Arg Ala Val Ser Ala Asp Gln
                2725                2730                2735

ttg gcg ggc ggc gcg gac ccg gcc gtc cgc gat gcg ctg ttc cgc gtg   8256
Leu Ala Gly Gly Ala Asp Pro Ala Val Arg Asp Ala Leu Phe Arg Val
            2740                2745                2750

cag tgg aac ccc gta cgc ctg ccc ccg gcc ggg gcc gcg gtg acc gtg   8304
Gln Trp Asn Pro Val Arg Leu Pro Pro Ala Gly Ala Ala Val Thr Val
        2755                2760                2765

gcg acg ctc ggc tcc ctt gcc ggc gca ccg ttc gac ggc tac ccg gac   8352
Ala Thr Leu Gly Ser Leu Ala Gly Ala Pro Phe Asp Gly Tyr Pro Asp
    2770                2775                2780

ctg gcg tcc ctg gcc cgg tcc ggt cgt gtg gcg ggt gcg gtg ctg gta   8400
Leu Ala Ser Leu Ala Arg Ser Gly Arg Val Ala Gly Ala Val Leu Val
2785                2790                2795                2800

ccg gtg gaa gcc ggt gcc ggc gag gtg gtg gcg gac gat gtc gtg ggg   8448
Pro Val Glu Ala Gly Ala Gly Glu Val Val Ala Asp Asp Val Val Gly
                2805                2810                2815
```

```
gcg acg cac gca acg gcc gcc cgg gcg ctg gac ctg gcc cgg tcg tgg   8496
Ala Thr His Ala Thr Ala Ala Arg Ala Leu Asp Leu Ala Arg Ser Trp
            2820                2825                2830

ctg gcc gat gac cgg ttc gcg gcc tcg cgc ctg gtg ttc gtg acg cgt   8544
Leu Ala Asp Asp Arg Phe Ala Ala Ser Arg Leu Val Phe Val Thr Arg
       2835                2840                2845

ggc gcg gtg tcc ggt gcg gat ctc gcg ggt gcg gcg gtg tgg ggt ctg   8592
Gly Ala Val Ser Gly Ala Asp Leu Ala Gly Ala Ala Val Trp Gly Leu
   2850                2855                2860

gtg cgg tcg gcg ctg tcg gag cac ccg ggc cgc ttc ggt ctg gtg gat   8640
Val Arg Ser Ala Leu Ser Glu His Pro Gly Arg Phe Gly Leu Val Asp
2865                2870                2875                2880

ctg gat gac gat gcc gaa ctg gcg ctg gtg cca cgg gtg ttg gcg tcg   8688
Leu Asp Asp Asp Ala Glu Leu Ala Leu Val Pro Arg Val Leu Ala Ser
               2885                2890                2895

gat gag ccg cag ctg ctg gtg cgc ggt ggt gag gtg ctg gcg gcg cgg   8736
Asp Glu Pro Gln Leu Leu Val Arg Gly Gly Glu Val Leu Ala Ala Arg
            2900                2905                2910

ctg gcc cgg gcg cag tcc tcg cac gcg gtg acc tgg gat ccg tcc ggc   8784
Leu Ala Arg Ala Gln Ser Ser His Ala Val Thr Trp Asp Pro Ser Gly
       2915                2920                2925

acg gtg ctc gtc acc ggt ggc acg ggt ggt ctg ggc cgt gtg atg gca   8832
Thr Val Leu Val Thr Gly Gly Thr Gly Gly Leu Gly Arg Val Met Ala
   2930                2935                2940

cgt cac ttg gtg gtg gaa cac ggg gta cgg aac ctg ctg ctg gtc agc   8880
Arg His Leu Val Val Glu His Gly Val Arg Asn Leu Leu Leu Val Ser
2945                2950                2955                2960

cgc cgt ggg ccc gcc gcc gaa ggt gcc gaa gag ctg gtg acg gag ctc   8928
Arg Arg Gly Pro Ala Ala Glu Gly Ala Glu Glu Leu Val Thr Glu Leu
               2965                2970                2975

cgg cac agc ggt gcc gaa gtg gcc gtc gaa gcc tgt gat gtc acc gac   8976
Arg His Ser Gly Ala Glu Val Ala Val Glu Ala Cys Asp Val Thr Asp
            2980                2985                2990

gcg gcc gcc gtg gcc gac ctg gtg gcc cgg cac cgg atc agc gct gtg   9024
Ala Ala Ala Val Ala Asp Leu Val Ala Arg His Arg Ile Ser Ala Val
       2995                3000                3005

gtg cat acg gcc ggt gtc ctg gat gac ggt gtg gtg gag tcg ctg aca   9072
Val His Thr Ala Gly Val Leu Asp Asp Gly Val Val Glu Ser Leu Thr
   3010                3015                3020

ccg gag cgg ctg tcg gcg gtg ttg cgt ccg aag gtg gat gcg gcc tgg   9120
Pro Glu Arg Leu Ser Ala Val Leu Arg Pro Lys Val Asp Ala Ala Trp
3025                3030                3035                3040

aac ctg cac gag gcg acc agg gat ctg gac ctg gac gcg ttc gtg gtc   9168
Asn Leu His Glu Ala Thr Arg Asp Leu Asp Leu Asp Ala Phe Val Val
               3045                3050                3055

ttc tcc tca gtg gca ggc acg atc ggg agc ccc ggt cag gcc aac tac   9216
Phe Ser Ser Val Ala Gly Thr Ile Gly Ser Pro Gly Gln Ala Asn Tyr
            3060                3065                3070

gcg gcg ggc aac gcc ttc ctg gat gcc ctg gcc cac cac cgt cgg gcg   9264
```

```
Ala Ala Gly Asn Ala Phe Leu Asp Ala Leu Ala His His Arg Arg Ala
       3075                3080                3085

gcg ggt ctt ccg gcg gcg tcg ctg gca tgg ggc ccc tgg tcc cgg gac   9312
Ala Gly Leu Pro Ala Ala Ser Leu Ala Trp Gly Pro Trp Ser Arg Asp
   3090                3095                3100

ggc ggc atg acc ggc acc ctg acc gac gtc gac tcc agc gca tcg ccc   9360
Gly Gly Met Thr Gly Thr Leu Thr Asp Val Asp Ser Ser Ala Ser Pro
3105                3110                3115                3120

ggc agg cat gcc cga act cac ccc cgc aca ggg cgt ggc ctc ttc gac   9408
Gly Arg His Ala Arg Thr His Pro Arg Thr Gly Arg Gly Leu Phe Asp
               3125                3130                3135

gcc gcg ctg gcg gcc ggt gac gcc cac ctg ctc ccc gta cgc ttc gac   9456
Ala Ala Leu Ala Ala Gly Asp Ala His Leu Leu Pro Val Arg Phe Asp
           3140                3145                3150

tgg gcg tcc ctg cgc gcc cag ggc gag gtg cca ccg ctg ttg cgc ggc   9504
Trp Ala Ser Leu Arg Ala Gln Gly Glu Val Pro Pro Leu Leu Arg Gly
       3155                3160                3165

ctg atc agg acc cgt gcc cgg cgc tcg gcg gtc ggc ggc tcg gcc gcg   9552
Leu Ile Arg Thr Arg Ala Arg Arg Ser Ala Val Gly Gly Ser Ala Ala
   3170                3175                3180

gca gcc ggc ctg gtg gga cgc ctg agc gga cgg gga acg gtg gag cgg   9600
Ala Ala Gly Leu Val Gly Arg Leu Ser Gly Arg Gly Thr Val Glu Arg
3185                3190                3195                3200

cgc gag gtg ctc ctg gac ctg gta cgg gcc cag atc gcg gtc gtc ctg   9648
Arg Glu Val Leu Leu Asp Leu Val Arg Ala Gln Ile Ala Val Val Leu
               3205                3210                3215

ggc cac gcg aac ccg gag acg atc gag tcc acc cgt gtc ttc cag gac   9696
Gly His Ala Asn Pro Glu Thr Ile Glu Ser Thr Arg Val Phe Gln Asp
           3220                3225                3230

ctc ggc ttc gac tcc ctg acc gcg gtc gaa ctc cgc aac cgc ctc aac   9744
Leu Gly Phe Asp Ser Leu Thr Ala Val Glu Leu Arg Asn Arg Leu Asn
       3235                3240                3245

aac gcg acc ggc ctg cgc ctt tcg gcc acc gcc gtc ttc gac tac ccc   9792
Asn Ala Thr Gly Leu Arg Leu Ser Ala Thr Ala Val Phe Asp Tyr Pro
   3250                3255                3260

acg gcg gac gcg ctc gtc gac ttc ctg ctg gac gag ctg ttc ggc gcg   9840
Thr Ala Asp Ala Leu Val Asp Phe Leu Leu Asp Glu Leu Phe Gly Ala
3265                3270                3275                3280

cag gag gag gcc gag ctg ccg gcg ccg gtg ccg tca ccg gcg ggg gcc   9888
Gln Glu Glu Ala Glu Leu Pro Ala Pro Val Pro Ser Pro Ala Gly Ala
               3285                3290                3295

gcc gac gac ccg gtc gtg atc gtc ggc atg agc tgc cgc tac ccg ggc   9936
Ala Asp Asp Pro Val Val Ile Val Gly Met Ser Cys Arg Tyr Pro Gly
           3300                3305                3310

ggc gtc ggc tcg ccc gag gac ctg tgg cgc ctg gtg tcg gag ggc gtg   9984
Gly Val Gly Ser Pro Glu Asp Leu Trp Arg Leu Val Ser Glu Gly Val
       3315                3320                3325

gac gcg gtg tcc gac ttc ccc acc gac cgt gga tgg gac gtg gag agc   10032
Asp Ala Val Ser Asp Phe Pro Thr Asp Arg Gly Trp Asp Val Glu Ser
```

```
   3330                3335                3340

ctc tac agc ccc gac ccc gag gcg ctc ggc acc tcg tac acc cgc tcc   10080
Leu Tyr Ser Pro Asp Pro Glu Ala Leu Gly Thr Ser Tyr Thr Arg Ser
3345                3350                3355                3360

ggt gga ttc ctc cac gag gcg gcg gag ttc gac ccc gat ttc ttc ggg   10128
Gly Gly Phe Leu His Glu Ala Ala Glu Phe Asp Pro Asp Phe Phe Gly
                3365                3370                3375

atg agc ccg cgc gag gcg ctg gcg acc gac gcc cag cag cgg ctg ctg   10176
Met Ser Pro Arg Glu Ala Leu Ala Thr Asp Ala Gln Gln Arg Leu Leu
            3380                3385                3390

ctg gag acg acc tgg gag gcc atc gag cgc acg ggc atc gac ccg gcg   10224
Leu Glu Thr Thr Trp Glu Ala Ile Glu Arg Thr Gly Ile Asp Pro Ala
        3395                3400                3405

tcg ctg cgg ggc agc cgt acg ggc gtc ttc gcg ggc gtg atg tac acc   10272
Ser Leu Arg Gly Ser Arg Thr Gly Val Phe Ala Gly Val Met Tyr Thr
    3410                3415                3420

gac tac ggc gac ctc ctc gtc ggc gac cag ttc gag ggc tac cgc agc   10320
Asp Tyr Gly Asp Leu Leu Val Gly Asp Gln Phe Glu Gly Tyr Arg Ser
3425                3430                3435                3440

aac ggc agc gcg gcc agc atc gcc tcc ggc cgg gtc tcg tac acc ttc   10368
Asn Gly Ser Ala Ala Ser Ile Ala Ser Gly Arg Val Ser Tyr Thr Phe
                3445                3450                3455

ggt ttc gag ggt ccg gcg gtc acg gtg gac acg gca tgc tcg tcg tcc   10416
Gly Phe Glu Gly Pro Ala Val Thr Val Asp Thr Ala Cys Ser Ser Ser
            3460                3465                3470

ctg gtc gcc ctg cac tgg gcg gcg cag tcg ctg cgc tcg ggc gag tgc   10464
Leu Val Ala Leu His Trp Ala Ala Gln Ser Leu Arg Ser Gly Glu Cys
        3475                3480                3485

tcg ctc gcg gtc gcg ggc ggt gtg acg gtg atg tcc aca ccg acg acg   10512
Ser Leu Ala Val Ala Gly Gly Val Thr Val Met Ser Thr Pro Thr Thr
    3490                3495                3500

ttc gtc gag ttc tcg cgg caa cgc gga ctg tcg gcg gac ggc cgc tgc   10560
Phe Val Glu Phe Ser Arg Gln Arg Gly Leu Ser Ala Asp Gly Arg Cys
3505                3510                3515                3520

aag gcg ttc gcc gat gcg gcc gac ggc gtc ggc tgg ggc gag ggc gtc   10608
Lys Ala Phe Ala Asp Ala Ala Asp Gly Val Gly Trp Gly Glu Gly Val
                3525                3530                3535

ggc atg ctc gta ctg gag cgt ctg tcg gac gcg cgc cgc aac ggg cac   10656
Gly Met Leu Val Leu Glu Arg Leu Ser Asp Ala Arg Arg Asn Gly His
            3540                3545                3550

cgg gtg ctc gcg gtg gtg cgc ggc agt gcg gtg aac cag gac ggt gcg   10704
Arg Val Leu Ala Val Val Arg Gly Ser Ala Val Asn Gln Asp Gly Ala
        3555                3560                3565

tcc aat ggt ctg acg gcg ccg aac ggc ccc gcc cag cag cgg gtg atc   10752
Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ala Gln Gln Arg Val Ile
    3570                3575                3580

cgg cag gcg ctg gcg agt gcg ggg ctg tcg gcg gcg gat gtg gac gcg   10800
Arg Gln Ala Leu Ala Ser Ala Gly Leu Ser Ala Ala Asp Val Asp Ala
3585                3590                3595                3600
```

```
gtg gag gcg cac ggt acg ggt acg acg ctg ggc gat ccg atc gag gcc   10848
Val Glu Ala His Gly Thr Gly Thr Thr Leu Gly Asp Pro Ile Glu Ala
                3605                3610                3615

cag gcg ctg ctc gcc acg tat ggc cag gag cga cct gag gac cgg ccg   10896
Gln Ala Leu Leu Ala Thr Tyr Gly Gln Glu Arg Pro Glu Asp Arg Pro
            3620                3625                3630

ttg ctg ctg ggg tcg gtc aaa tcc aac atc ggt cat gcg cag gcg gct   10944
Leu Leu Leu Gly Ser Val Lys Ser Asn Ile Gly His Ala Gln Ala Ala
        3635                3640                3645

tcg ggt gtg gcg ggt gtc atc aag atg gtg ctg gcg atg cgg cac ggt   10992
Ser Gly Val Ala Gly Val Ile Lys Met Val Leu Ala Met Arg His Gly
    3650                3655                3660

gtg ctg cct cgg acg ctg cat gtg gat gaa ccg tcg tcg cat gtc gac   11040
Val Leu Pro Arg Thr Leu His Val Asp Glu Pro Ser Ser His Val Asp
3665                3670                3675                3680

tgg agt gcc ggt gcc gtc gag ctg ctg acc tcc gag gcc gag tgg ccg   11088
Trp Ser Ala Gly Ala Val Glu Leu Leu Thr Ser Glu Ala Glu Trp Pro
                3685                3690                3695

cag ggc gag ggg ccg cgc cgc gcg ggc gtc tcc tcc ttc ggc gtc agc   11136
Gln Gly Glu Gly Pro Arg Arg Ala Gly Val Ser Ser Phe Gly Val Ser
            3700                3705                3710

ggg acg aac gcg cat gtg atc ctg gag cag ccc gga ccg gac gcg gcc   11184
Gly Thr Asn Ala His Val Ile Leu Glu Gln Pro Gly Pro Asp Ala Ala
        3715                3720                3725

gac gcc gca ccg gac gcc acg gtg acc gat ccc ggc gcg ctg gca tgg   11232
Asp Ala Ala Pro Asp Ala Thr Val Thr Asp Pro Gly Ala Leu Ala Trp
    3730                3735                3740

gtg ctc tcc gca cgg aac gaa gcg gcc ctg cgc tgc cag gcg gcg cgc   11280
Val Leu Ser Ala Arg Asn Glu Ala Ala Leu Arg Cys Gln Ala Ala Arg
3745                3750                3755                3760

ctg ctg tcc ctg gtc gcc ggc agt gac gcg ctg tgc gcg cgg gac atc   11328
Leu Leu Ser Leu Val Ala Gly Ser Asp Ala Leu Cys Ala Arg Asp Ile
                3765                3770                3775

ggc cac tcg ctg gtg acc ggg cgg tcg agc ttc gcc cac cgt gcg gtg   11376
Gly His Ser Leu Val Thr Gly Arg Ser Ser Phe Ala His Arg Ala Val
            3780                3785                3790

gtg tgg ggc cag gac cgc gac gca ctg gtg cgt gcc ctg tcc gca ctc   11424
Val Trp Gly Gln Asp Arg Asp Ala Leu Val Arg Ala Leu Ser Ala Leu
        3795                3800                3805

gcg gtg ggc gag gcc gac gcc ggt ctg gcg gag ggc gcg tcc ggc gcg   11472
Ala Val Gly Glu Ala Asp Ala Gly Leu Ala Glu Gly Ala Ser Gly Ala
    3810                3815                3820

ggg agg acg gcc ttc ctg ttc tcg ggc cag gga tca caa cgg ctg gga   11520
Gly Arg Thr Ala Phe Leu Phe Ser Gly Gln Gly Ser Gln Arg Leu Gly
3825                3830                3835                3840

atg gga tgg gag ttg tac gct cgc tac ccg gtg ttc gcg gac gca ttc   11568
Met Gly Trp Glu Leu Tyr Ala Arg Tyr Pro Val Phe Ala Asp Ala Phe
                3845                3850                3855
```

```
gac gcc gtg tgc gcg gcc ttg gac gag cac ctg gag cgc ccc ctg cgg   11616
Asp Ala Val Cys Ala Ala Leu Asp Glu His Leu Glu Arg Pro Leu Arg
            3860                3865                3870

gac gtg gtc tgg ggc gag gac gcg gag ctg ctg aac cag acc gcg tac   11664
Asp Val Val Trp Gly Glu Asp Ala Glu Leu Leu Asn Gln Thr Ala Tyr
        3875                3880                3885

gcc cag gcc ggg ctg ttc gcg atc gag gtg gcg ctg tac cgg ctg gcg   11712
Ala Gln Ala Gly Leu Phe Ala Ile Glu Val Ala Leu Tyr Arg Leu Ala
    3890                3895                3900

gaa tcg tgg ggc atg cgc ccg gac ttc gtg gcg ggg cat tcg atc ggt   11760
Glu Ser Trp Gly Met Arg Pro Asp Phe Val Ala Gly His Ser Ile Gly
3905                3910                3915                3920

gag gtc gcc gcg gcc cat gtg tcg ggt gtc ttc tcg ctc ccg gat gcc   11808
Glu Val Ala Ala Ala His Val Ser Gly Val Phe Ser Leu Pro Asp Ala
                3925                3930                3935

tgt gcg ctg gtg gcg gcc cga ggc cga ctg atg cag caa ctg ccc tcc   11856
Cys Ala Leu Val Ala Ala Arg Gly Arg Leu Met Gln Gln Leu Pro Ser
            3940                3945                3950

ggc ggc gcg atg atg gcg atc cgg gcg acc gag gac gag gtc ctt ccg   11904
Gly Gly Ala Met Met Ala Ile Arg Ala Thr Glu Asp Glu Val Leu Pro
        3955                3960                3965

cat ctg gcg gaa ggc gtc tcg ctc gcg gcg gtc aat ggc ccg tcg tcg   11952
His Leu Ala Glu Gly Val Ser Leu Ala Ala Val Asn Gly Pro Ser Ser
    3970                3975                3980

gtc gtg atc tcg ggc gcc gag gac gcg gtg ctg gcc atc gcg gcg cac   12000
Val Val Ile Ser Gly Ala Glu Asp Ala Val Leu Ala Ile Ala Ala His
3985                3990                3995                4000

ttc gcg ggg gag ggg cgc aaa acc acc cga ctg cgg gtc tcg cat gcc   12048
Phe Ala Gly Glu Gly Arg Lys Thr Thr Arg Leu Arg Val Ser His Ala
                4005                4010                4015

ttc cac tcg ccg ctc atg gaa ccg atg ctg gag gaa ttc cgc gcg gtg   12096
Phe His Ser Pro Leu Met Glu Pro Met Leu Glu Glu Phe Arg Ala Val
            4020                4025                4030

gtg aca cgg ctg tcc ttc ggc acg ccg acg atc ccc gtc gtc tcc aac   12144
Val Thr Arg Leu Ser Phe Gly Thr Pro Thr Ile Pro Val Val Ser Asn
        4035                4040                4045

ctg acg ggc cgc ctc gcc gaa ccc gaa cag ctc gcg cac gcc gac tac   12192
Leu Thr Gly Arg Leu Ala Glu Pro Glu Gln Leu Ala His Ala Asp Tyr
    4050                4055                4060

tgg gtc cgg cac gtc cgc gag gca gtg cgc ttc gcg gac ggg ata cag   12240
Trp Val Arg His Val Arg Glu Ala Val Arg Phe Ala Asp Gly Ile Gln
4065                4070                4075                4080

gcg ctg cgg gcg gaa ggg gtg acg cgg ttc ctg gag ctc ggc ccg gac   12288
Ala Leu Arg Ala Glu Gly Val Thr Arg Phe Leu Glu Leu Gly Pro Asp
                4085                4090                4095

ggt gtg ctg tcg gcg atg gcc cgc gag tcg gca tcg gac gac gcc gtg   12336
Gly Val Leu Ser Ala Met Ala Arg Glu Ser Ala Ser Asp Asp Ala Val
            4100                4105                4110

ctc gcg ccc gta ctg cgc agg gac cgg ccc gag gag acg gcg ctg ctg   12384
```

```
Leu Ala Pro Val Leu Arg Arg Asp Arg Pro Glu Glu Thr Ala Leu Leu
        4115                4120                4125

ggc gcc ctg gcg cag ctg tac gtc cgg ggt gcg cac gtg gac tgg acg   12432
Gly Ala Leu Ala Gln Leu Tyr Val Arg Gly Ala His Val Asp Trp Thr
    4130                4135                4140

gtg ccg ttc gcc ggt tcg ggt gcg cgc tgg gcg gat ctg ccg acg tac   12480
Val Pro Phe Ala Gly Ser Gly Ala Arg Trp Ala Asp Leu Pro Thr Tyr
4145                4150                4155                4160

gcg ttc cag cac gag cgg ttc tgg ccg tcg ggc ggt gtg gca cgt ccg   12528
Ala Phe Gln His Glu Arg Phe Trp Pro Ser Gly Gly Val Ala Arg Pro
                4165                4170                4175

ggc gat gtg cgg tcc gcg ggc ctg ggc tcg gcc ggg cat ccg ctg ctg   12576
Gly Asp Val Arg Ser Ala Gly Leu Gly Ser Ala Gly His Pro Leu Leu
            4180                4185                4190

ggc gcg gcg gtg gaa ctg gcg ggc tcg ggc ggc ctg ttg ttc acg ggc   12624
Gly Ala Ala Val Glu Leu Ala Gly Ser Gly Gly Leu Leu Phe Thr Gly
        4195                4200                4205

cgg ctg tcg gtg tcc tcg cac ccg tgg ctg gcg gac cat gtg gtg ctg   12672
Arg Leu Ser Val Ser Ser His Pro Trp Leu Ala Asp His Val Val Leu
    4210                4215                4220

ggc tcc gtc ctc gtg ccc ggc acc gcg ctg gtg gaa ctg gtg ctg cgg   12720
Gly Ser Val Leu Val Pro Gly Thr Ala Leu Val Glu Leu Val Leu Arg
4225                4230                4235                4240

gcg gcc gac gag gcc ggc tgt gac ctc ctg gag gag ctg acg ctc gcc   12768
Ala Ala Asp Glu Ala Gly Cys Asp Leu Leu Glu Glu Leu Thr Leu Ala
                4245                4250                4255

gca ccg ctg gtg ctg ccc gcc tcg ggc gcc gcg gtc cag gtt cag gta   12816
Ala Pro Leu Val Leu Pro Ala Ser Gly Ala Ala Val Gln Val Gln Val
            4260                4265                4270

gcg gtg ggc gag ccc gat gag gcg ggc cgc cgg ccg gtc tcg gtc cat   12864
Ala Val Gly Glu Pro Asp Glu Ala Gly Arg Arg Pro Val Ser Val His
        4275                4280                4285

gca cgt gag ggc gag ggc cca tgg acg ctg cac gcc agt ggt gcg gtg   12912
Ala Arg Glu Gly Glu Gly Pro Trp Thr Leu His Ala Ser Gly Ala Val
    4290                4295                4300

acc tcg ggc gcc gaa gtg ccc ccc ttc gac gcc acc gta tgg ccg ccc   12960
Thr Ser Gly Ala Glu Val Pro Pro Phe Asp Ala Thr Val Trp Pro Pro
4305                4310                4315                4320

aag ggc gcg gag ccc gtg gac gtg gcg gac tgc tac gac gta ctc gcc   13008
Lys Gly Ala Glu Pro Val Asp Val Ala Asp Cys Tyr Asp Val Leu Ala
                4325                4330                4335

gat gcc ggg ctc acc tac ggc ccg gcc ttc cac ggc ctg caa gcg gcc   13056
Asp Ala Gly Leu Thr Tyr Gly Pro Ala Phe His Gly Leu Gln Ala Ala
            4340                4345                4350

tgg aag ctc ggt ggg gac gtc tac gcc gag gcg aag ctc ccc gag agc   13104
Trp Lys Leu Gly Gly Asp Val Tyr Ala Glu Ala Lys Leu Pro Glu Ser
        4355                4360                4365

acc gac ggc gac gca tac ggt ctg cac ccc gcg ctc ttc gac gcc gcg   13152
Thr Asp Gly Asp Ala Tyr Gly Leu His Pro Ala Leu Phe Asp Ala Ala
```

```
   4370                4375                4380

ctg cac gcg tcg gcg ctg ggc ggc gcg gaa gcg ggc gga gtc ccg ttc   13200
Leu His Ala Ser Ala Leu Gly Gly Ala Glu Ala Gly Gly Val Pro Phe
4385                4390                4395                4400

tcc tgg gcc gga gtg tcg ctg cac gcg acc ggc gcc tcg cac ctc cgc   13248
Ser Trp Ala Gly Val Ser Leu His Ala Thr Gly Ala Ser His Leu Arg
                4405                4410                4415

gtc cgc atc cgc gaa gcg ggc ggc gcg ctg tcg gtc gcg atc gcg gac   13296
Val Arg Ile Arg Glu Ala Gly Gly Ala Leu Ser Val Ala Ile Ala Asp
            4420                4425                4430

acg tcc ggc gcg ccg gtc gcc tcg gtg gag tcg ctg gtg ata cgt ccg   13344
Thr Ser Gly Ala Pro Val Ala Ser Val Glu Ser Leu Val Ile Arg Pro
        4435                4440                4445

ctc tcg gcc ggg cag gtg cag gcc gcc gac cgt gac gcc ctc ttc aag   13392
Leu Ser Ala Gly Gln Val Gln Ala Ala Asp Arg Asp Ala Leu Phe Lys
    4450                4455                4460

gcc gac tgg gtc ccc gta ccg ctc acg gac gaa cgc gtc gag ccg ggc   13440
Ala Asp Trp Val Pro Val Pro Leu Thr Asp Glu Arg Val Glu Pro Gly
4465                4470                4475                4480

acc ggc ccg gag ggc gag ccg ctg cgg acg tac gcg gat ctg gat tcc   13488
Thr Gly Pro Glu Gly Glu Pro Leu Arg Thr Tyr Ala Asp Leu Asp Ser
                4485                4490                4495

ctg gag ggc gcg gcc gtg ccc ggg acg gtc ctg gtc gcg ccg cct tcc   13536
Leu Glu Gly Ala Ala Val Pro Gly Thr Val Leu Val Ala Pro Pro Ser
            4500                4505                4510

ggc gct gcc ggg acg gtg gag tcc gta cac gcc gcg acc gtc tgg gcg   13584
Gly Ala Ala Gly Thr Val Glu Ser Val His Ala Ala Thr Val Trp Ala
        4515                4520                4525

ctg gag atg gtg cag gcg tgg ctg gcc gac gac cgg ttc gcc acc tcg   13632
Leu Glu Met Val Gln Ala Trp Leu Ala Asp Asp Arg Phe Ala Thr Ser
    4530                4535                4540

cga ctg gtg ttc gtc acc cgc ggc gcg gcc ttc ggc gcg gat ctt gcg   13680
Arg Leu Val Phe Val Thr Arg Gly Ala Ala Phe Gly Ala Asp Leu Ala
4545                4550                4555                4560

gcg gcc gcc gtc cgg ggc ctg gtg cgc tcg gca cag tcg gag aac ccg   13728
Ala Ala Ala Val Arg Gly Leu Val Arg Ser Ala Gln Ser Glu Asn Pro
                4565                4570                4575

ggc cgc ttc ggc ctg gtg gac atg gac ggc gac gcc gat acg acc gta   13776
Gly Arg Phe Gly Leu Val Asp Met Asp Gly Asp Ala Asp Thr Thr Val
            4580                4585                4590

ccg gcg caa gcg ctc gcg acc gac gag ccc gaa ctg ctg gtg cgt ggt   13824
Pro Ala Gln Ala Leu Ala Thr Asp Glu Pro Glu Leu Leu Val Arg Gly
        4595                4600                4605

ggt gag gtg ctg gcg gcc cgg ctg gtc cgg gcg cag tcc tcg cac acg   13872
Gly Glu Val Leu Ala Ala Arg Leu Val Arg Ala Gln Ser Ser His Thr
    4610                4615                4620

gtg acg tgg gat ccg tcc ggt acg gtc ctg atc acc ggc ggg acc ggt   13920
Val Thr Trp Asp Pro Ser Gly Thr Val Leu Ile Thr Gly Gly Thr Gly
4625                4630                4635                4640
```

```
ggg ctg ggc cgt agt gtc gcc cgg cac ttg gtg agc gag cac ggg gtg   13968
Gly Leu Gly Arg Ser Val Ala Arg His Leu Val Ser Glu His Gly Val
                4645                4650                4655

cgc agt ctg ctg ctg gtc agc cgc cgt ggt ccc gcg gcc gag ggt gcc   14016
Arg Ser Leu Leu Leu Val Ser Arg Arg Gly Pro Ala Ala Glu Gly Ala
            4660                4665                4670

ggg gag ttg gtg gcc gaa ctc agg ggc agt ggc gcc gag gtg gtc atc   14064
Gly Glu Leu Val Ala Glu Leu Arg Gly Ser Gly Ala Glu Val Val Ile
        4675                4680                4685

gag gct tgt gat gtg acc gat gcg gtg gcg gtg gcc gat ctg gtg gct   14112
Glu Ala Cys Asp Val Thr Asp Ala Val Ala Val Ala Asp Leu Val Ala
    4690                4695                4700

cgg cat cgg atc agt gct gtg gtg cat acg gcc ggt gtt ctg gat gac   14160
Arg His Arg Ile Ser Ala Val Val His Thr Ala Gly Val Leu Asp Asp
4705                4710                4715                4720

ggt gtg gtg gag tcg ctg acg ccg gag cgg ctt gcg gtg gtg ttg cgt   14208
Gly Val Val Glu Ser Leu Thr Pro Glu Arg Leu Ala Val Val Leu Arg
                4725                4730                4735

ccg aag gtg gat gcg gcc tgg aac ctg cac gag gcg acc agg ggt ctg   14256
Pro Lys Val Asp Ala Ala Trp Asn Leu His Glu Ala Thr Arg Gly Leu
            4740                4745                4750

gat ctg gat gcg ttt gtg gtg ttc tcg tcc gtg gca ggc act ttc ggc   14304
Asp Leu Asp Ala Phe Val Val Phe Ser Ser Val Ala Gly Thr Phe Gly
        4755                4760                4765

agt gcg ggt cag gcc aat tac gcg gcg ggt aat gct ttc ctg gac gcg   14352
Ser Ala Gly Gln Ala Asn Tyr Ala Ala Gly Asn Ala Phe Leu Asp Ala
    4770                4775                4780

ctg gcg tat cac cgt cgg gcg gtg ggt ctg ccg gcg gtg tcg ctg gcg   14400
Leu Ala Tyr His Arg Arg Ala Val Gly Leu Pro Ala Val Ser Leu Ala
4785                4790                4795                4800

tgg ggc cct tgg tcg cag gac ggt ggt atg acc ggc acc ttg agc gac   14448
Trp Gly Pro Trp Ser Gln Asp Gly Gly Met Thr Gly Thr Leu Ser Asp
                4805                4810                4815

gcc gat gtc cag cgc atc gcc cgg cag ggc atg ccg ccg ctg acc gtc   14496
Ala Asp Val Gln Arg Ile Ala Arg Gln Gly Met Pro Pro Leu Thr Val
            4820                4825                4830

gag gag ggt ctg gcc ctc ttc gac gcc gcg ctc ggc agc gcc gaa ccc   14544
Glu Glu Gly Leu Ala Leu Phe Asp Ala Ala Leu Gly Ser Ala Glu Pro
        4835                4840                4845

atg gca ctc ccg gtc cgc ctg gac ctg gcg gcg ctg cgg gca caa ggc   14592
Met Ala Leu Pro Val Arg Leu Asp Leu Ala Ala Leu Arg Ala Gln Gly
    4850                4855                4860

gag ccc cag cca ctg ctg cgc ggc ctc atc cgg acg agg acc cgc cgg   14640
Glu Pro Gln Pro Leu Leu Arg Gly Leu Ile Arg Thr Arg Thr Arg Arg
4865                4870                4875                4880

tcc ggc gcc gcc gcg gca tcc ggc atc gcg cag cgc ctt gcc ggg ctg   14688
Ser Gly Ala Ala Ala Ala Ser Gly Ile Ala Gln Arg Leu Ala Gly Leu
                4885                4890                4895
```

```
tcc acg gcg gag cgg cgc gag gcg ctg ctc gat gtc gta cgg gcc cag   14736
Ser Thr Ala Glu Arg Arg Glu Ala Leu Leu Asp Val Val Arg Ala Gln
            4900                4905                4910

atc gcg acg gtc ctg ggc cac gcc ggc ccg gaa acg atc gcc cct gac   14784
Ile Ala Thr Val Leu Gly His Ala Gly Pro Glu Thr Ile Ala Pro Asp
        4915                4920                4925

cgg gcc ttc cag gac ctc ggc ctc gac tcc ctg acg gcg atc gaa ctc   14832
Arg Ala Phe Gln Asp Leu Gly Leu Asp Ser Leu Thr Ala Ile Glu Leu
    4930                4935                4940

cgt aac ctg ctc ggc aag gcc acc ggg ctg cgg ctc ccg gca acg acc   14880
Arg Asn Leu Leu Gly Lys Ala Thr Gly Leu Arg Leu Pro Ala Thr Thr
4945                4950                4955                4960

gtg ttc gac tac ccg acg gtg gat gcc ctg gcc gcc cac ctc ttg gac   14928
Val Phe Asp Tyr Pro Thr Val Asp Ala Leu Ala Ala His Leu Leu Asp
                4965                4970                4975

gaa ctg ttc ggc gcg gag acg ggg acc gcg acg gag acg ccc ctc ccg   14976
Glu Leu Phe Gly Ala Glu Thr Gly Thr Ala Thr Glu Thr Pro Leu Pro
            4980                4985                4990

gtg ccc ggc ctg ccg tcc ctg gcg gac gat ccg gtc gtg atc gtc ggc   15024
Val Pro Gly Leu Pro Ser Leu Ala Asp Asp Pro Val Val Ile Val Gly
        4995                5000                5005

atg agc tgc cgc ttc ccc ggc ggc gtc gcc tcg ccg gag gac ctg tgg   15072
Met Ser Cys Arg Phe Pro Gly Gly Val Ala Ser Pro Glu Asp Leu Trp
    5010                5015                5020

cgc ctg gtg gcg gac ggc gtg gac gcc gtc tcc gcc ttc ccg acc gac   15120
Arg Leu Val Ala Asp Gly Val Asp Ala Val Ser Ala Phe Pro Thr Asp
5025                5030                5035                5040

cgg ggc tgg gag atc gac gac acc tac gac ccc gag cgg gag ggc gcc   15168
Arg Gly Trp Glu Ile Asp Asp Thr Tyr Asp Pro Glu Arg Glu Gly Ala
                5045                5050                5055

atc gcc acc cgt tcc ggt gga ttc ctc cac gac gcg gcg gag ttc gac   15216
Ile Ala Thr Arg Ser Gly Gly Phe Leu His Asp Ala Ala Glu Phe Asp
            5060                5065                5070

ccc gag ttc ttc ggg atg agc ccg cgc gag gcc ctg acc acc gac gcc   15264
Pro Glu Phe Phe Gly Met Ser Pro Arg Glu Ala Leu Thr Thr Asp Ala
        5075                5080                5085

cag cag cgg ctg ttg ctg gag acg acc tgg gag gcg ctg gag cgc gcc   15312
Gln Gln Arg Leu Leu Leu Glu Thr Thr Trp Glu Ala Leu Glu Arg Ala
    5090                5095                5100

ggt atg gac ccg gcc acg ctc cgc ggc agc cgc acg ggt gtc ttc gcc   15360
Gly Met Asp Pro Ala Thr Leu Arg Gly Ser Arg Thr Gly Val Phe Ala
5105                5110                5115                5120

ggc gtg atg tac cac gac tac tcg acg ctg ctc tcc ggg cgc gag ttc   15408
Gly Val Met Tyr His Asp Tyr Ser Thr Leu Leu Ser Gly Arg Glu Phe
                5125                5130                5135

gag ggc tac cag ggc agc ggc agc gca ggc agt gtg gcc tcg ggc cgg   15456
Glu Gly Tyr Gln Gly Ser Gly Ser Ala Gly Ser Val Ala Ser Gly Arg
            5140                5145                5150

gtc tcg tac acc ttc ggt ttc gag ggt ccg gcg gtc acg gtg gac acg   15504
```

```
Val Ser Tyr Thr Phe Gly Phe Glu Gly Pro Ala Val Thr Val Asp Thr
      5155                5160                5165

gcg tgc tcg tcg tcc ctg gtc gcc ctg cac ctg gca gca cag tcg ctg   15552
Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala Ala Gln Ser Leu
   5170                5175                5180

cgc tcg ggc gag tgc tcg ctg gcg ctc gcg ggc ggt gtg acg gtg atg   15600
Arg Ser Gly Glu Cys Ser Leu Ala Leu Ala Gly Gly Val Thr Val Met
5185                5190                5195                5200

tcc aca ccg ctg acc ttc gtg gag ttc tcc cgc cag ggc gga ctg tcg   15648
Ser Thr Pro Leu Thr Phe Val Glu Phe Ser Arg Gln Gly Gly Leu Ser
               5205                5210                5215

gcg gac ggc cgc tgc aag gcg ttc gcc gat gcg gcc gac ggc gtc ggc   15696
Ala Asp Gly Arg Cys Lys Ala Phe Ala Asp Ala Ala Asp Gly Val Gly
            5220                5225                5230

tgg gcc gaa ggc gcc gga atc ctg gtg ctg gag cgt ctg tcg gac gcc   15744
Trp Ala Glu Gly Ala Gly Ile Leu Val Leu Glu Arg Leu Ser Asp Ala
      5235                5240                5245

cgc cgc aac ggg cac cgc atc ctc gcg acg gtg cgc ggc agt gcg gtg   15792
Arg Arg Asn Gly His Arg Ile Leu Ala Thr Val Arg Gly Ser Ala Val
   5250                5255                5260

aac cag gac ggt gcg tcc aat ggt ctg acg gcg ccg aac ggt ccc gcc   15840
Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ala
5265                5270                5275                5280

cag cag cgg gtg atc cgg cag gcg ctg gcg agt gcg ggg ctg tcg gcg   15888
Gln Gln Arg Val Ile Arg Gln Ala Leu Ala Ser Ala Gly Leu Ser Ala
               5285                5290                5295

gcg gat gtg gac gcg gtg gag gcg cac ggt acg ggt acg acg ctg ggc   15936
Ala Asp Val Asp Ala Val Glu Ala His Gly Thr Gly Thr Thr Leu Gly
            5300                5305                5310

gat ccg atc gag gcc cag gcg ctg ctc gcg acg tat ggc cag gag cgg   15984
Asp Pro Ile Glu Ala Gln Ala Leu Leu Ala Thr Tyr Gly Gln Glu Arg
      5315                5320                5325

ccg gag gac cgg ccg ttg ctg ctc ggc tcc gtg aag tcc aac atc ggt   16032
Pro Glu Asp Arg Pro Leu Leu Leu Gly Ser Val Lys Ser Asn Ile Gly
   5330                5335                5340

cac gcg caa gcg gct tcg ggt gtt gcc ggt gtc atc aag atg gtg ctg   16080
His Ala Gln Ala Ala Ser Gly Val Ala Gly Val Ile Lys Met Val Leu
5345                5350                5355                5360

gcg atg cgg cac ggt gtg ctg cct cgg acg ctg cat gtc gac gag ccg   16128
Ala Met Arg His Gly Val Leu Pro Arg Thr Leu His Val Asp Glu Pro
               5365                5370                5375

tcg tcg cat gtc gac tgg agc gcc ggt gcc gtc gag ctg ctg acc tcc   16176
Ser Ser His Val Asp Trp Ser Ala Gly Ala Val Glu Leu Leu Thr Ser
            5380                5385                5390

gag gcc gag tgg ccg cag ggc gag ggg ccg cgc cgc gcg ggc gtc tcc   16224
Glu Ala Glu Trp Pro Gln Gly Glu Gly Pro Arg Arg Ala Gly Val Ser
      5395                5400                5405

tcc ttc ggc atc agt ggg acg aac gcg cat gtg atc ctg gag cag ccc   16272
Ser Phe Gly Ile Ser Gly Thr Asn Ala His Val Ile Leu Glu Gln Pro
```

```
5410                5415                5420

gaa ccg gtc gcg gcg gaa acg gaa tcg atc acg ccc gac acc gca ccg   16320
Glu Pro Val Ala Ala Glu Thr Glu Ser Ile Thr Pro Asp Thr Ala Pro
5425                5430                5435                5440

gac gcc gcc gag gac gag gcg gcc gat tcc ggg acg ccg gtg ccg gca   16368
Asp Ala Ala Glu Asp Glu Ala Ala Asp Ser Gly Thr Pro Val Pro Ala
                5445                5450                5455

ctg ctg tcc ggc agg agc gca tcg gcg ctg cgg gcc cag gca gca cga   16416
Leu Leu Ser Gly Arg Ser Ala Ser Ala Leu Arg Ala Gln Ala Ala Arg
            5460                5465                5470

ctg ctg tcc cga ctc gac ggc gat ccg ggg ccg cgg atc act gac gtc   16464
Leu Leu Ser Arg Leu Asp Gly Asp Pro Gly Pro Arg Ile Thr Asp Val
        5475                5480                5485

gcc tac tcc ctc gcg acc ggc cgt tcg gcc ttc ccg cac cgc gcg gtg   16512
Ala Tyr Ser Leu Ala Thr Gly Arg Ser Ala Phe Pro His Arg Ala Val
    5490                5495                5500

atc ctc gcc gcg aac cga gcg gac ctg ctg cac tcg ctg tcc gcc ctg   16560
Ile Leu Ala Ala Asn Arg Ala Asp Leu Leu His Ser Leu Ser Ala Leu
5505                5510                5515                5520

gcc gag ggc cac acc gag gcg ccg gcc gta gtc gca cag gac cga gcc   16608
Ala Glu Gly His Thr Glu Ala Pro Ala Val Val Ala Gln Asp Arg Ala
                5525                5530                5535

cgc tcg ggc aag ctg gcc ttc ctg ttc tcg ggg cag gga tcg caa cgc   16656
Arg Ser Gly Lys Leu Ala Phe Leu Phe Ser Gly Gln Gly Ser Gln Arg
            5540                5545                5550

ctg ggc atg gga cgg gag ttg tac ggt cgc tac ccg gcg ttc gcc gag   16704
Leu Gly Met Gly Arg Glu Leu Tyr Gly Arg Tyr Pro Ala Phe Ala Glu
        5555                5560                5565

gcc ctc gac gcg gtg tgc gcc gcc ctg gac gcc cac ctg gac cgt ccc   16752
Ala Leu Asp Ala Val Cys Ala Ala Leu Asp Ala His Leu Asp Arg Pro
    5570                5575                5580

ctg cgg gac gtc atc tgg ggc gag gac gcg gaa ctg ctg aac cgg acc   16800
Leu Arg Asp Val Ile Trp Gly Glu Asp Ala Glu Leu Leu Asn Arg Thr
5585                5590                5595                5600

ggg tac gcc cag aca ggg ctg ttc gcc atc gag gtg gcc ctg ttc cgc   16848
Gly Tyr Ala Gln Thr Gly Leu Phe Ala Ile Glu Val Ala Leu Phe Arg
                5605                5610                5615

ctg ctg gag tcg tgg ggc gta cgc ccg gac cac ctg ctg ggg cac tcc   16896
Leu Leu Glu Ser Trp Gly Val Arg Pro Asp His Leu Leu Gly His Ser
            5620                5625                5630

atc gga gaa atc gcc gcg gcc cat gtg gcg ggc gtc ctc tcc ctc ccg   16944
Ile Gly Glu Ile Ala Ala Ala His Val Ala Gly Val Leu Ser Leu Pro
        5635                5640                5645

gac gcc tgt gcg ctg gtg gcg gcc cga ggt cgg ctg atg cag caa ctg   16992
Asp Ala Cys Ala Leu Val Ala Ala Arg Gly Arg Leu Met Gln Gln Leu
    5650                5655                5660

ccg tcc ggc ggc gcg atg atg gcg atc cgg gcg acc gag gac gag gtc   17040
Pro Ser Gly Gly Ala Met Met Ala Ile Arg Ala Thr Glu Asp Glu Val
5665                5670                5675                5680
```

```
ctt ccg cat ctg gcg gaa ggc gtc tcg ctc gcg gcg gtc aac ggg ccg   17088
Leu Pro His Leu Ala Glu Gly Val Ser Leu Ala Ala Val Asn Gly Pro
                5685                5690                5695

tcg tcg gtc gtg gtc tcc ggc gcc gag gac gag gta ctc gcc ctc gcg   17136
Ser Ser Val Val Val Ser Gly Ala Glu Asp Glu Val Leu Ala Leu Ala
            5700                5705                5710

gcg cac ttc gag gaa gag gga cgc aag acc acc cga ctg cgg gtc tcg   17184
Ala His Phe Glu Glu Glu Gly Arg Lys Thr Thr Arg Leu Arg Val Ser
        5715                5720                5725

cac gcc ttc cac tcc ccg ctc atg gaa ccg atg ctg gcc gac ttc cgg   17232
His Ala Phe His Ser Pro Leu Met Glu Pro Met Leu Ala Asp Phe Arg
    5730                5735                5740

gcc gtc gcc gac ggc atg acc tac gcc gcg ccg cgc atc ccc gtg gtc   17280
Ala Val Ala Asp Gly Met Thr Tyr Ala Ala Pro Arg Ile Pro Val Val
5745                5750                5755                5760

tcc aac gtc acc ggc cgg ccc gcc acc gcg gaa gag ctg tgc tgc gcc   17328
Ser Asn Val Thr Gly Arg Pro Ala Thr Ala Glu Glu Leu Cys Cys Ala
                5765                5770                5775

gag tac tgg gtc ggc cac gta cgc gag gcc gta cgg ttc gcc gac ggg   17376
Glu Tyr Trp Val Gly His Val Arg Glu Ala Val Arg Phe Ala Asp Gly
            5780                5785                5790

gtc ggc gcg ctc cgc gag cag ggt gtg acg acg ttc ctg gaa ctc ggc   17424
Val Gly Ala Leu Arg Glu Gln Gly Val Thr Thr Phe Leu Glu Leu Gly
        5795                5800                5805

ccc gac ggc tct ctc tcc gcg ctc gcc gcc gaa tcc gcc gcc gac gac   17472
Pro Asp Gly Ser Leu Ser Ala Leu Ala Ala Glu Ser Ala Ala Asp Asp
    5810                5815                5820

tcc gta ctg gcc ccc gta ctg cgc aag aac cgc ccc gag gca ccg gca   17520
Ser Val Leu Ala Pro Val Leu Arg Lys Asn Arg Pro Glu Ala Pro Ala
5825                5830                5835                5840

ctg ctc acg gcc ctg gca cga ctg cac gcc cag ggc acg ccg gtc gac   17568
Leu Leu Thr Ala Leu Ala Arg Leu His Ala Gln Gly Thr Pro Val Asp
                5845                5850                5855

tgg tcc gcc gcc ttc gcc ggt acg ggt gcg cgg tgg gtg gac ctg ccg   17616
Trp Ser Ala Ala Phe Ala Gly Thr Gly Ala Arg Trp Val Asp Leu Pro
            5860                5865                5870

acg tac gca ttc cag cac gag cgg ttc tgg ccg tcg ggc ggg gcg gcg   17664
Thr Tyr Ala Phe Gln His Glu Arg Phe Trp Pro Ser Gly Gly Ala Ala
        5875                5880                5885

cgc gca ggc gat gtg cgg tcc gcg ggc ctg ggc tcg gcc ggg cac ccg   17712
Arg Ala Gly Asp Val Arg Ser Ala Gly Leu Gly Ser Ala Gly His Pro
    5890                5895                5900

ctg ctg ggt gct gcg gtg gaa ctg gcg ggc tcc ggc ggg cgg ttg ctc   17760
Leu Leu Gly Ala Ala Val Glu Leu Ala Gly Ser Gly Gly Arg Leu Leu
5905                5910                5915                5920

acc ggg cgg ctg tcc ctg tcc tcg cac ccg tgg ctg gcg gat cac gtg   17808
Thr Gly Arg Leu Ser Leu Ser Ser His Pro Trp Leu Ala Asp His Val
                5925                5930                5935
```

```
gtg ctg ggc tcc gta ctg gtg ccc ggc acg gcg ctc atg gaa ctg gtg   17856
Val Leu Gly Ser Val Leu Val Pro Gly Thr Ala Leu Met Glu Leu Val
            5940                5945                5950

ctg cgg gcg gcc gac gag gtg gac tgc gcc gcg gtg gac gaa ctc acg   17904
Leu Arg Ala Ala Asp Glu Val Asp Cys Ala Ala Val Asp Glu Leu Thr
        5955                5960                5965

ctc gcc gcg cca ctg gtc ctg ccc gcc tcg ggc gcc gcg atc cag gta   17952
Leu Ala Ala Pro Leu Val Leu Pro Ala Ser Gly Ala Ala Ile Gln Val
    5970                5975                5980

cag gta tgg gtg ggc gag ccc gat gag gcg ggc cgc cgg ccg gtc tcg   18000
Gln Val Trp Val Gly Glu Pro Asp Glu Ala Gly Arg Arg Pro Val Ser
5985                5990                5995                6000

gtc cat gca cgc gag ggc gag ggc cca tgg acg ctg cac gcc gac ggc   18048
Val His Ala Arg Glu Gly Glu Gly Pro Trp Thr Leu His Ala Asp Gly
                6005                6010                6015

gcc ctg gcc ccg gcg gcc gag acg gtg ccg ttc gat acc gcg ata tgg   18096
Ala Leu Ala Pro Ala Ala Glu Thr Val Pro Phe Asp Thr Ala Ile Trp
            6020                6025                6030

ccc ccg cag ggt gcc gag cac ctg gac gcg gcg ggc tgt tac gag cgg   18144
Pro Pro Gln Gly Ala Glu His Leu Asp Ala Ala Gly Cys Tyr Glu Arg
        6035                6040                6045

ttc gcg gac gcc gga ttc gcg tac ggc ccg gtg ttc cag ggc ctg cgg   18192
Phe Ala Asp Ala Gly Phe Ala Tyr Gly Pro Val Phe Gln Gly Leu Arg
    6050                6055                6060

gcg gcc tgg aag ctc ggc gag gac atc tac gcc gag gtc gca ctc ccc   18240
Ala Ala Trp Lys Leu Gly Glu Asp Ile Tyr Ala Glu Val Ala Leu Pro
6065                6070                6075                6080

gaa ggc acg gac ggc aac gcc tac ggc ctg cac ccc gca ctc ttc gac   18288
Glu Gly Thr Asp Gly Asn Ala Tyr Gly Leu His Pro Ala Leu Phe Asp
                6085                6090                6095

gcc gcg ctg cac gca gcg ctc ctg ggc ggc gag gga acg gac gaa gcc   18336
Ala Ala Leu His Ala Ala Leu Leu Gly Gly Glu Gly Thr Asp Glu Ala
            6100                6105                6110

gcg gtc ccc ttc tcc tgg aac ggg gtg acg ctc cac gcc acc ggc gct   18384
Ala Val Pro Phe Ser Trp Asn Gly Val Thr Leu His Ala Thr Gly Ala
        6115                6120                6125

tcc cgg gtg agg gta cgc atc cgt ccc acc gaa ggc ggt acg tcg ata   18432
Ser Arg Val Arg Val Arg Ile Arg Pro Thr Glu Gly Gly Thr Ser Ile
    6130                6135                6140

gcc ctc gtg gac acc gcc ggt gcg ccg gtc gcc tcg gtg cga tcc ctg   18480
Ala Leu Val Asp Thr Ala Gly Ala Pro Val Ala Ser Val Arg Ser Leu
6145                6150                6155                6160

acc gca cgt ccg atc acc gcc ggg cag ttg cag acc ggt gac cgc gat   18528
Thr Ala Arg Pro Ile Thr Ala Gly Gln Leu Gln Thr Gly Asp Arg Asp
                6165                6170                6175

tcc ctt ttc cag gtc gac tgg acc acc ctc cac ctc acg gac gag cgc   18576
Ser Leu Phe Gln Val Asp Trp Thr Thr Leu His Leu Thr Asp Glu Arg
            6180                6185                6190

gcg aac tcc ctc gcc ctg ctc ggc aag gac acc gag ggc atc ctc gac   18624
```

```
Ala Asn Ser Leu Ala Leu Leu Gly Lys Asp Thr Glu Gly Ile Leu Asp
       6195                6200                6205

aca ctc tcc ctc cag ccc cac gcg gac ctc gac gac ctc gcg gcg acg   18672
Thr Leu Ser Leu Gln Pro His Ala Asp Leu Asp Asp Leu Ala Ala Thr
   6210                6215                6220

ggc gtc cac gac acc gtg ctc gcc ccg ctg ccc acc cgg acc gcc gga   18720
Gly Val His Asp Thr Val Leu Ala Pro Leu Pro Thr Arg Thr Ala Gly
6225                6230                6235                6240

acg gtg gaa tcc gtc cat gcc gcc acg aca ggg gca ctg gcc ctg atc   18768
Thr Val Glu Ser Val His Ala Ala Thr Thr Gly Ala Leu Ala Leu Ile
                6245                6250                6255

cgg tcc tgg ctg gcc gac gac cgg ttc gcc gcc tcg cgc ctg gtg ttc   18816
Arg Ser Trp Leu Ala Asp Asp Arg Phe Ala Ala Ser Arg Leu Val Phe
            6260                6265                6270

gtg acg cgt ggc gcg gtg tcc ggc acg gat ctc gcg ggt gcg tcg gtg   18864
Val Thr Arg Gly Ala Val Ser Gly Thr Asp Leu Ala Gly Ala Ser Val
       6275                6280                6285

tgg ggc ctg gtg cgg tcg gcg ttg ttg gag cac ccg ggc cgc ttc ggt   18912
Trp Gly Leu Val Arg Ser Ala Leu Leu Glu His Pro Gly Arg Phe Gly
   6290                6295                6300

ctg gtg gac gtg gac gtg gac caa gac gct gaa gtg ccg ctt gtg cca   18960
Leu Val Asp Val Asp Val Asp Gln Asp Ala Glu Val Pro Leu Val Pro
6305                6310                6315                6320

agg gcg ttg gcg tcg gat gaa ccg cag gtg ttg gtg cgt ggt ggt gag   19008
Arg Ala Leu Ala Ser Asp Glu Pro Gln Val Leu Val Arg Gly Gly Glu
                6325                6330                6335

gtg ctg gcg gcc cgg ctg gtc cgg gcg cag tcc tcg gac acg gtg acg   19056
Val Leu Ala Ala Arg Leu Val Arg Ala Gln Ser Ser Asp Thr Val Thr
            6340                6345                6350

tgg gat ccg tcc ggt acg gtc ctg atc acc ggc ggg acc ggt ggg ctg   19104
Trp Asp Pro Ser Gly Thr Val Leu Ile Thr Gly Gly Thr Gly Gly Leu
       6355                6360                6365

ggt cgt agt gtc gcc cgg cac ttg gtg agc gag cac ggg gtg cgc agt   19152
Gly Arg Ser Val Ala Arg His Leu Val Ser Glu His Gly Val Arg Ser
   6370                6375                6380

ctg ctg ctg gtc agc cgc cgt ggt ccc gcc gcc gag ggt gtc gat gca   19200
Leu Leu Leu Val Ser Arg Arg Gly Pro Ala Ala Glu Gly Val Asp Ala
6385                6390                6395                6400

ctc gtt gcc gaa ctt gcc gag tgc ggc gcg cag gtc acc gtc gag gct   19248
Leu Val Ala Glu Leu Ala Glu Cys Gly Ala Gln Val Thr Val Glu Ala
                6405                6410                6415

tgt gat gtg act gac gcg gtg gcg gtg gcc gat ctg gtg gct cgg cat   19296
Cys Asp Val Thr Asp Ala Val Ala Val Ala Asp Leu Val Ala Arg His
            6420                6425                6430

cgg atc agt gct gtg gtg cat acg gcc ggt gtt ctg gat gac ggt gtg   19344
Arg Ile Ser Ala Val Val His Thr Ala Gly Val Leu Asp Asp Gly Val
       6435                6440                6445

gtg gag tcg ctg acg ccg gag cgg ctg tcg gcg gtg ctg cgt ccg aag   19392
Val Glu Ser Leu Thr Pro Glu Arg Leu Ser Ala Val Leu Arg Pro Lys
```

```
   6450                6455                6460

gtg gat gcg gcc tgg aac ctg cac gag gcg acc agg ggt ctg gat ctg   19440
Val Asp Ala Ala Trp Asn Leu His Glu Ala Thr Arg Gly Leu Asp Leu
6465                6470                6475                6480

gat gcg ttt gtg gtg ttc tcg tcc gtg gca ggc acc ttc ggc agt gcg   19488
Asp Ala Phe Val Val Phe Ser Ser Val Ala Gly Thr Phe Gly Ser Ala
                6485                6490                6495

ggt cag gcc aat tac gcg gcg ggt aat gct ttc ctg gac gcg ctg gcg   19536
Gly Gln Ala Asn Tyr Ala Ala Gly Asn Ala Phe Leu Asp Ala Leu Ala
            6500                6505                6510

tat cac cgt cgg gcg gtg ggt ttg ccg gcg gtg tcg ctg gcg tgg ggc   19584
Tyr His Arg Arg Ala Val Gly Leu Pro Ala Val Ser Leu Ala Trp Gly
        6515                6520                6525

cct tgg tcg cag gac ggt ggt atg acc ggc acc ttg agc gac gcc gat   19632
Pro Trp Ser Gln Asp Gly Gly Met Thr Gly Thr Leu Ser Asp Ala Asp
    6530                6535                6540

gtc cag cgc atc gcc cgg cag ggc atg ccg ccg ctg acc gtc gag gag   19680
Val Gln Arg Ile Ala Arg Gln Gly Met Pro Pro Leu Thr Val Glu Glu
6545                6550                6555                6560

ggt ctg gcc ctc ttc gac gcc gcg ctc ggc agc gcc gaa ccc atg gca   19728
Gly Leu Ala Leu Phe Asp Ala Ala Leu Gly Ser Ala Glu Pro Met Ala
                6565                6570                6575

ctc ccc gtc cgc ctg gac ctc gcg gcc cta cgg gca caa ggc gag ccc   19776
Leu Pro Val Arg Leu Asp Leu Ala Ala Leu Arg Ala Gln Gly Glu Pro
            6580                6585                6590

cag cca ctg ctg cgc ggc ctc atc cgg acc ccg ggt cga cgc acg gcg   19824
Gln Pro Leu Leu Arg Gly Leu Ile Arg Thr Pro Gly Arg Arg Thr Ala
        6595                6600                6605

gcg gcc gcg acg gag ggc gac acc gct gcc gcc ttc gcc ggg cgc ctg   19872
Ala Ala Ala Thr Glu Gly Asp Thr Ala Ala Ala Phe Ala Gly Arg Leu
    6610                6615                6620

acc ggg ctg tcg gcg gca gaa gga cgc gag gtc gta ctg ggc gcc gta   19920
Thr Gly Leu Ser Ala Ala Glu Gly Arg Glu Val Val Leu Gly Ala Val
6625                6630                6635                6640

cgc agc cag atc gcg ggg gtc ctc gga cac gcc gaa gcc acg gaa atc   19968
Arg Ser Gln Ile Ala Gly Val Leu Gly His Ala Glu Ala Thr Glu Ile
                6645                6650                6655

gac cag gac cgc gcc ttc ctg gac ctc gga ttc gac tcc ctc acc gcg   20016
Asp Gln Asp Arg Ala Phe Leu Asp Leu Gly Phe Asp Ser Leu Thr Ala
            6660                6665                6670

gtc gaa ctc cgc aac cgc ctg ggc gcc gtc acc gga atc cgc ctg ccg   20064
Val Glu Leu Arg Asn Arg Leu Gly Ala Val Thr Gly Ile Arg Leu Pro
        6675                6680                6685

gcg acc ctg ctc ttc gac tac ccg acg ccg gca gaa ctc gtc gcc cac   20112
Ala Thr Leu Leu Phe Asp Tyr Pro Thr Pro Ala Glu Leu Val Ala His
    6690                6695                6700

ctc cat gcc cgg atc gca ccg gag ccg acc gtc ggc ccg gag gcg ctc   20160
Leu His Ala Arg Ile Ala Pro Glu Pro Thr Val Gly Pro Glu Ala Leu
6705                6710                6715                6720
```

```
ctg ggc gaa ctc gaa agg atg gag aag tcc ttc ggc gga ctc gac ctc   20208
Leu Gly Glu Leu Glu Arg Met Glu Lys Ser Phe Gly Gly Leu Asp Leu
                6725                6730                6735

acg gag gag atg cac gaa cag ata gcc ggc cgt ctg gaa gtc ctc cgg   20256
Thr Glu Glu Met His Glu Gln Ile Ala Gly Arg Leu Glu Val Leu Arg
            6740                6745                6750

gcc aag tgg gac gcc ctg cgg gac acg gca gcg gca gcc ggg cac gac   20304
Ala Lys Trp Asp Ala Leu Arg Asp Thr Ala Ala Ala Ala Gly His Asp
        6755                6760                6765

ggt tcc ccg tcc gac gag gac ttc gac ttc gag tcc gcc tcc gac gac   20352
Gly Ser Pro Ser Asp Glu Asp Phe Asp Phe Glu Ser Ala Ser Asp Asp
    6770                6775                6780

gag gtc ttc gac ctc ctc gac aac gaa ctc ggc ctg tcc tga           20394
Glu Val Phe Asp Leu Leu Asp Asn Glu Leu Gly Leu Ser
6785                6790                6795
```

<210> 4
<211> 6798
<212> PRT
<213> Streptomyces natalensis

<400> 4

```
Met Ser Asn Glu Glu Lys Leu Arg Glu Tyr Leu Lys Arg Ala Ile Ala
  1               5                  10                  15

Asp Leu His Glu Thr Arg Gln Gln Leu Asp Glu Thr Glu Ala Lys Gln
             20                  25                  30

Arg Glu Pro Leu Ala Ile Val Ser Met Ala Cys Arg Phe Pro Gly Gly
         35                  40                  45

Val Arg Ser Pro Glu Glu Leu Trp Glu Leu Leu Arg Asp Gly Val Asp
     50                  55                  60

Ala Val Ser Ser Phe Pro Arg Asn Arg Gly Trp Asp Leu Asp Ala Leu
 65                  70                  75                  80

Tyr His Ser Asp Pro Ala His Gln Gly Thr Ser Tyr Ala Arg Glu Gly
                 85                  90                  95

Gly Phe Leu His Asp Ala Gly Glu Phe Asp Pro Gly Phe Phe Gly Ile
            100                 105                 110

Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg Leu Leu Leu
        115                 120                 125

Glu Thr Ala Trp Glu Ala Val Glu Arg Ala Gly Ile Asp Pro Glu Ser
    130                 135                 140

Leu Ala Gly Ser Arg Thr Gly Val Phe Val Gly Thr Gly His Gly Gly
145                 150                 155                 160

Tyr Asp Ala Glu Gly Arg Arg Arg Ala Asp Glu Val Gly Gly His Leu
                165                 170                 175

Leu Thr Gly Asn His Ile Ser Ile Ala Ser Gly Arg Ile Ser Tyr Val
            180                 185                 190

Leu Gly Leu Glu Gly Pro Ala Leu Thr Val Asp Thr Ala Cys Ser Ser
```

195 200 205

Ser Leu Val Ala Leu His Leu Ala Met His Ala Leu Arg Arg Asp Glu
210 215 220

Cys Ala Met Ala Leu Val Gly Gly Ala Thr Val Met Ser Thr Pro Gln
225 230 235 240

Met Phe Val Glu Phe Ser Arg Gln Arg Gly Leu Ala Pro Asp Gly Arg
245 250 255

Cys Lys Pro Phe Ala Ala Ala Ala Asp Gly Thr Gly Trp Ser Glu Gly
260 265 270

Val Gly Leu Leu Leu Val Glu Arg Leu Ser Asp Ala Val Arg Asn Gly
275 280 285

Tyr Pro Val Leu Ala Val Leu Lys Gly Ser Ala Val Asn Gln Asp Gly
290 295 300

Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ser Gln Gln Arg Val
305 310 315 320

Ile Arg Gln Ala Leu Thr Gly Ala Gly Leu Ala Ala Ser Asp Ile Asp
325 330 335

Ala Val Glu Ala His Gly Thr Gly Thr Thr Leu Gly Asp Pro Val Glu
340 345 350

Ala His Ala Leu Leu Ala Thr Tyr Gly Gln Gln Arg Ala Ala Asp Arg
355 360 365

Pro Cys Gly Leu Gly Ser Met Lys Ser Asn Ile Gly His Thr Gln Ala
370 375 380

Ala Ala Gly Ile Ala Gly Val Met Lys Met Val Leu Ala Met Arg His
385 390 395 400

Gly His Leu Pro Arg Thr Leu His Leu Asp Glu Pro Thr Gly His Val
405 410 415

Asp Trp Ser Glu Gly Asn Ala Arg Leu Leu Ala Glu Pro Glu Pro Trp
420 425 430

Pro Ser Ala Gly Arg Pro Arg Arg Ala Ala Val Ser Ser Phe Gly Ile
435 440 445

Ser Gly Thr Asn Ala His Val Ile Leu Glu Gln Ala Pro Ala His Glu
450 455 460

Ala Glu Pro Ala Pro Glu Pro Ala Ala Arg Pro Gly Ala Leu Pro Trp
465 470 475 480

Ile Leu Ser Ala Arg Thr Glu Ala Gly Leu Arg Ala Gln Ala Asp Arg
485 490 495

Leu Gly Arg His Leu Arg Asp Arg Ala Asp Leu Glu Pro Ala Ala Val
500 505 510

Ala His Ala Leu Ala Asp Thr Arg Thr Leu Met Glu His Arg Ala Val
515 520 525

Val Val Ala Gly Asp Arg Glu Glu Phe Leu Arg Gly Leu Asp Ala Leu
530 535 540

```
Ala Ala Gly Arg Thr Ala Asn Gly Leu Val Ser Gly Val Ala Val Lys
545                 550                 555                 560

Ala Ala Ser Ala Phe Leu Phe Ala Gly Gln Gly Ser Gln Arg Pro Gly
                565                 570                 575

Met Gly Arg Glu Leu His Ala Ala His Pro Val Phe Ala Thr Ala Phe
            580                 585                 590

Asp Ala Val Cys Ala Glu Leu Asp Pro His Leu Asp Arg Pro Leu Arg
        595                 600                 605

Asp Ile Val Phe Ala Glu Glu Gly Ser Ala Glu Ala Ala Leu Leu Asp
    610                 615                 620

Gln Thr Ala Tyr Thr Gln Ala Ala Leu Phe Ala Leu Glu Thr Ala Leu
625                 630                 635                 640

Phe Arg Leu Val Glu Ser Trp Gly Val Ala Pro Arg Phe Val Ala Gly
                645                 650                 655

His Ser Ile Gly Glu Leu Thr Ala Ala His Val Ser Gly Val Leu Thr
            660                 665                 670

Leu His Asp Ala Ala Arg Leu Val Ala Ala Arg Gly Thr Leu Met Gln
        675                 680                 685

Ala Leu Pro Ala Gly Gly Ala Met Val Ala Val Gln Ala Thr Glu Asp
    690                 695                 700

Glu Ile Arg Glu Arg Leu Ala Gly His Glu Asp His Val Ala Leu Ala
705                 710                 715                 720

Ala Ala Asn Gly Pro Asp Ser Thr Val Ile Ser Gly Asp Glu Gln Ala
                725                 730                 735

Val Thr Glu Ile Ala Ala His Trp Glu Ala Gln Gly Arg Arg Thr Lys
            740                 745                 750

Arg Leu Arg Val Ser His Ala Phe His Ser Pro His Met Asp Asp Met
        755                 760                 765

Leu Glu Asp Phe Arg Arg Val Ala Arg Gly Leu Thr Phe His Ala Pro
    770                 775                 780

Arg Ile Pro Val Val Ser Thr Val Thr Gly Ala Leu Ala Thr Glu Asp
785                 790                 795                 800

Glu Leu Arg Ser Pro Asp Tyr Trp Val Arg Gln Val Arg Glu Thr Val
                805                 810                 815

Arg Phe Cys Ala Ala Val Arg Thr Leu Glu Ala Glu Gly Val Thr Thr
            820                 825                 830

Phe Val Glu Ile Gly Thr Gly Gly Val Leu Thr Pro Met Val Gln Asp
        835                 840                 845

Cys Leu Thr Thr Leu Glu Glu Pro Val Leu Val Pro Leu Leu Arg Thr
    850                 855                 860

Gly Arg Pro Glu Thr Val Ala Leu Thr Glu Gly Val Ala Thr Ala Phe
865                 870                 875                 880

Val His Gly Val Pro Val Asp Arg Ser Ala Phe Pro Gly Ala Pro Gly
                885                 890                 895
```

Thr Ser Arg Ala Asp Leu Pro Thr Tyr Ala Phe Gln Arg Gln Trp Tyr
900 905 910

Trp Leu Asp Pro Ala Asp His Asp Glu Gly Glu Ala Ala Ala Ala Glu
915 920 925

Ala Gly Glu Ala Gly Phe Trp Ala Ala Val Glu Arg Glu Asp Leu Gln
930 935 940

Glu Leu Ser Ala Val Leu Ala Ile Asp Gly Ser Glu Ala Asp Ser Leu
945 950 955 960

Gly Ser Phe Leu Pro Thr Leu Ser Ser Trp Arg Arg Gln Arg Arg Thr
965 970 975

Gln Ala Ala Ala Asp Arg Phe Ser Tyr Arg Thr His Trp Ala Pro Arg
980 985 990

Thr Ala Ser Gly Gly Pro Thr Ala Thr Gly His Trp Leu Val Val Leu
995 1000 1005

Pro Glu Gly Gly Thr Asp Asp Pro Trp Thr Ala Arg Leu Leu Asp Ala
1010 1015 1020

Leu Asn Asp Gln Gly Leu His Thr Asp Val Arg Glu Leu Pro Ala Asp
025 1030 1035 1040

His Glu Pro Asp Ala Trp Gly Arg His Pro Val Asp Gly Val Leu Cys
1045 1050 1055

Leu Leu Ala Leu Asp Glu Arg Pro Thr Arg Ser Cys Pro Pro Tyr Arg
1060 1065 1070

Arg Gly Leu Ala Ala Thr Thr Asn Ala Ala Ala Arg Pro Glu Gly Ala
1075 1080 1085

Gly Ile Gln Ala Pro Leu Trp Cys Val Thr Arg Gly Ala Val Ala Val
1090 1095 1100

Asp Arg His Glu Ala Leu Lys Ser Pro Leu Gln Ala Gln Thr Trp Gly
105 1110 1115 1120

Leu Gly Arg Val Ala Ala Leu Glu Ser Pro Gln Ser Trp Gly Gly Leu
1125 1130 1135

Ile Asp Leu Pro Asp Asn Leu Asp Gly Arg Ala Val Ser Ala Leu Leu
1140 1145 1150

Ser Thr Leu Ala Gly Glu Glu Asp Gln Val Ala Val Arg Pro Ala Gly
1155 1160 1165

Val Phe Ala Arg Arg Leu Glu Arg Ile Thr Pro Gly Gly Asp Thr Gly
1170 1175 1180

Asp Arg Trp Ser Thr His Gly Thr Val Leu Val Thr Gly Gly Thr Gly
185 1190 1195 1200

Ala Leu Gly Ala His Leu Ala His Trp Leu Ala Asp Ala Gly Ala Glu
1205 1210 1215

His Leu Val Leu Thr Gly Arg Arg Gly Pro Gln Ala Pro Gly Ala Pro
1220 1225 1230

Glu Leu Ala Ala Ala Leu Thr Asp Arg Gly Val Lys Val Thr Leu Ala

1235 1240 1245

Ala Cys Asp Ala Ala Asp Arg Asp Ala Leu Ala Ala Val Leu Ala Asp
1250 1255 1260

Ile Pro Pro His Leu Pro Leu Thr Gly Val Val His Ala Ala Gly Val
265 1270 1275 1280

Leu Asp Asp Gly Val Leu Asp Ala Leu Thr Pro Glu Arg Phe Glu Thr
1285 1290 1295

Val Leu Arg Pro Lys Ala Arg Ala Ala Gln Asn Leu His Glu Leu Thr
1300 1305 1310

Gln Asp Leu Asp Leu Asp His Phe Val Leu Phe Ser Ser Ile Val Gly
1315 1320 1325

Val Leu Gly Asn Ala Gly Gln Ala Asn Tyr Ala Ala Ala Asn Ala Tyr
1330 1335 1340

Leu Asp Ala Leu Ala Glu His Arg Leu Ala Gln Gly Leu Pro Ala Thr
345 1350 1355 1360

Ser Val Ser Trp Gly Pro Gly Gln Ala Ala Ala Trp His Asp Ser Asp
1365 1370 1375

Ala Ala Asp Arg Met Ser Arg Asp Gly Leu Leu Pro Met Ala Ala Ala
1380 1385 1390

Pro Arg Arg Arg Pro Ala Pro Ala Leu Ala Gln Gly Met Thr Gln Val
1395 1400 1405

Thr Val Ala Asp Ile Asp Trp Ser Ala Tyr Ala Pro Ala Leu Thr Ala
1410 1415 1420

Val Arg Pro Ser Pro Leu Ile Gly Asp Leu Pro Glu Ala Arg Arg Ala
425 1430 1435 1440

Leu Gly Pro Ala Glu Gly Pro Arg Arg Glu Arg Ser Pro Leu Arg Asp
1445 1450 1455

Arg Ile Gly Ala Leu Pro Pro Ala Glu Gln Glu Lys Ala Phe Leu Thr
1460 1465 1470

Met Val Arg Glu Glu Ala Ala Arg Val Leu Gly His Pro Ser Pro Asp
1475 1480 1485

Thr Val Asp Ala Gln Arg Ala Phe Arg Glu Gln Gly Phe Asp Ser Leu
1490 1495 1500

Met Ala Val Asp Leu Arg Asn Arg Leu Ser Ala Ala Thr Gly Leu Arg
505 1510 1515 1520

Leu Pro Ala Thr Leu Leu Phe Asp His Pro Thr Pro Leu Ala Ala Ala
1525 1530 1535

Ala Cys Leu Arg Ser Glu Val Leu Gly Ala Ala Gly Pro Ala Thr Val
1540 1545 1550

Val Gln Ala Ser Thr Ala Ala Leu Asp Glu Pro Val Ala Ile Ile Gly
1555 1560 1565

Met Ala Cys Arg Phe Pro Gly Gly Val His Ser Pro Glu Ala Leu Trp
1570 1575 1580

Arg Leu Leu Ala Glu Gly Gly Asp Ala Ile Thr Pro Met Pro Ala Asp
585 1590 1595 1600

Arg Gly Trp Asp Leu Asp Arg Leu Tyr His Pro Asp Pro Asp His Gln
1605 1610 1615

Gly Thr Ser Tyr Ala Arg Gly Gly Gly Phe Leu Asp Gly Ala Ala Asp
1620 1625 1630

Phe Asp Ala Asp Phe Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met
1635 1640 1645

Asp Pro Gln Gln Arg Leu Leu Leu Glu Thr Trp Glu Val Leu Glu Gln
1650 1655 1660

Ala Gly Ile Asp Pro Glu Ser Leu Arg Gly Ser Ser Thr Gly Val Phe
665 1670 1675 1680

Ala Gly Thr Asn Thr Gln Asp Tyr Gly Thr Ala Leu Asp Ala Ala Gln
1685 1690 1695

Asp Glu Ala Gly Gly His Arg Leu Thr Gly Asn Ala Met Ser Val Val
1700 1705 1710

Ser Gly Arg Val Ser Tyr Thr Phe Gly Phe Glu Gly Pro Ala Leu Thr
1715 1720 1725

Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala Leu His Met Ala Ala
1730 1735 1740

Gln Ala Leu Arg Gln Gly Glu Cys Ser Leu Ala Val Ala Gly Gly Val
745 1750 1755 1760

Thr Val Met Ala Thr Pro Ser Ser Phe Val Glu Phe Ala Arg Gln Arg
1765 1770 1775

Gly Leu Ala Pro Asp Gly Arg Cys Lys Pro Phe Ala Ala Ala Ala Asp
1780 1785 1790

Gly Thr Gly Trp Ser Glu Gly Val Gly Leu Leu Leu Val Glu Arg Leu
1795 1800 1805

Ser Asp Ala Arg Arg Asn Gly His Gln Val Leu Ala Val Val Arg Gly
1810 1815 1820

Ser Ala Val Asn Gln Asp Gly Ala Ser Asn Gly Leu Ser Ala Pro Ser
825 1830 1835 1840

Gly Pro Ser Gln Gln Arg Val Ile Arg Gln Ala Leu Ala Asn Ala Arg
1845 1850 1855

Val Ala Ala Ser Glu Val Asp Ala Val Glu Ala His Gly Thr Gly Thr
1860 1865 1870

Thr Leu Gly Asp Pro Ile Glu Ala Gln Ala Leu Leu Ala Thr Tyr Gly
1875 1880 1885

Gln Glu Arg Pro Leu Leu Leu Gly Ala Val Lys Ser Asn Leu Gly His
1890 1895 1900

Thr Gln Ala Ala Ala Gly Val Ala Gly Val Met Lys Met Val Leu Ala
905 1910 1915 1920

Met Arg His Gly Met Leu Pro Arg Thr Leu His Val Asp Glu Pro Thr
1925 1930 1935

Gly His Val Asp Trp Thr Ala Gly Ala Val Glu Leu Leu Thr Glu His
1940 1945 1950

Thr Asp Trp Pro Glu Thr Gly His Pro Arg Arg Ala Ala Val Ser Ala
1955 1960 1965

Phe Gly Ile Ser Gly Thr Asn Ala His Val Val Leu Glu Leu Pro Ala
1970 1975 1980

Ala Glu Gln Pro Leu Val Glu Gln Pro Ser Ala Ala Glu Pro Asp Ala
985 1990 1995 2000

Pro Ala Thr Ala Pro Asp Arg Thr Pro Thr Ala Ser Asp Gly Thr Ala
2005 2010 2015

Pro Leu Leu Leu Ser Ala Lys Ser Glu Ser Ala Leu Arg Ala Gln Ala
2020 2025 2030

Ala Arg Leu His Ser His Leu Glu Arg Asp Pro Ala Leu Arg Leu Thr
2035 2040 2045

Asp Ala Ala Tyr Thr Leu Met Thr His Arg Thr Ala Phe Ala His Arg
2050 2055 2060

Ala Ala Val Arg Ala Ala Asp His Glu Ala Ala Leu Arg Ala Leu Thr
065 2070 2075 2080

Ala Leu Ala Ala Gly Glu Ala Asp Pro Ala Val Asp Thr Gly Thr Ala
2085 2090 2095

His Thr Gly Arg Asp Ala Val Leu Phe Ser Gly Gln Gly Ser Gln Arg
2100 2105 2110

Ile Gly Met Gly Arg Glu Leu Ser Gly Arg Tyr Pro Val Phe Ala Glu
2115 2120 2125

Ala Phe Asp Thr Val Cys Ala Ala Leu Asp Glu His Leu Asp Arg Pro
2130 2135 2140

Leu Arg Asp Val Val Arg Gly Glu Asp Glu Glu Leu Leu Asn Arg Thr
145 2150 2155 2160

Val Tyr Ala Gln Ala Gly Leu Phe Ala Ile Glu Val Ala Leu Phe Arg
2165 2170 2175

Leu Val Glu Ser Trp Gly Val Arg Pro His Tyr Val Ala Gly His Ser
2180 2185 2190

Val Gly Glu Ile Ala Ala Ala His Val Ala Gly Val Phe Ser Leu Ala
2195 2200 2205

Asp Ala Cys Ala Leu Val Ala Ala Arg Gly Arg Leu Met Gln Ala Leu
2210 2215 2220

Pro Ala Gly Gly Ala Met Ala Ala Ile Arg Ala Thr Glu Asp Glu Val
225 2230 2235 2240

Leu Pro His Leu Ala Asp Ser Val Ser Ile Ala Ala Val Asn Gly Pro
2245 2250 2255

Ser Ser Val Val Val Ser Gly Ala Glu His Ala Val Leu Ser Ile Ala
2260 2265 2270

Ala His Phe Glu Gly Ala Gly Arg Lys Thr Thr Arg Leu Arg Val Ser

2275 2280 2285

His Ala Phe His Ser Pro Leu Met Asp Pro Met Leu Ala Asp Phe Arg
2290 2295 2300

Ala Val Ala Glu Gly Leu Thr Tyr Gly Glu Pro Glu Leu Ala Val Val
305 2310 2315 2320

Ser Asn Val Thr Gly Gln Leu Ala Thr Pro Asp Gln Leu Arg Thr Pro
2325 2330 2335

Glu Tyr Trp Val Thr His Val Arg Ala Ala Val Arg Phe Ala Asp Gly
2340 2345 2350

Ile Arg Ala Leu Gly Ala Glu Gly Val Thr Arg Phe Leu Glu Leu Gly
2355 2360 2365

Pro Asp Gly Val Leu Ser Ala Leu Ala Arg Glu Ser Ala Pro Asp Asp
2370 2375 2380

Ala Val Cys Thr Pro Val Leu Arg Lys Asp Arg Ser Glu Ala Ala Thr
385 2390 2395 2400

Leu Leu Ala Ala Leu Thr His Leu His Val His Gly Thr Glu Ile Asp
2405 2410 2415

Trp Thr Ala Phe Leu Ala Gly Arg Asp Ala His Ala Val Asp Leu Pro
2420 2425 2430

Thr Tyr Ala Phe Gln His Gln Arg Phe Trp Pro Thr Pro Asp His Thr
2435 2440 2445

Arg Thr Gly Asp Leu Gly Ala Val Gly Leu Glu Ala Thr Gly His Pro
2450 2455 2460

Leu Leu Ser Ala Ala Val Glu Leu Pro Asp Gly Glu Gly Leu Leu Phe
465 2470 2475 2480

Thr Thr Arg Leu Ser Leu Gln Thr His Pro Trp Leu Ala Gly His Val
2485 2490 2495

Val Met Gly Ser Val Leu Leu Pro Gly Thr Ala Phe Ala Glu Leu Ala
2500 2505 2510

Leu Arg Ala Ala Asp Glu Val Gly Cys Asp Arg Val Asp Glu Leu Thr
2515 2520 2525

Leu Ala Ala Pro Leu Val Leu Pro Glu His Gly Gly Val Gln Leu Gln
2530 2535 2540

Leu Arg Val Gly Pro Ala Asp Ala Ser Gly Arg Arg Thr Leu Thr Ala
545 2550 2555 2560

Arg Ser Arg Ala Glu Gly Asp Gly Asp Arg Pro Trp Val Gln His Ala
2565 2570 2575

Thr Gly Val Leu Ala Glu Gly Glu Ser Thr Pro Glu Pro Gly Tyr Asp
2580 2585 2590

Phe His Thr Glu Ser Trp Pro Pro Ala Asp Ala Ala Pro Val Glu Leu
2595 2600 2605

Ser Gly Leu Tyr Pro Asp Phe Ala Ala His Gly Phe Asp Tyr Gly Pro
2610 2615 2620

```
His Phe Gln Gly Leu Arg Thr Ala Trp Arg Arg Gly Asp Glu Val Phe
625                 2630                2635                2640

Ala Glu Val Ala Leu Pro Ala Glu Ala Glu Gly Glu Ala Ser Ala Tyr
                2645                2650                2655

Gly Leu His Pro Ala Leu Leu Asp Ala Ala Leu His Val Val Ala Phe
            2660                2665                2670

Asn Gly Val Asp Arg Gly Val Val Pro Phe Ser Trp Glu Ser Val Ala
        2675                2680                2685

Leu His Ala Thr Gly Ala Ser Ala Val Arg Ile Arg Val Val Arg His
    2690                2695                2700

Ser Gly Asp Thr Val Ser Val Asp Val Ala Asp Thr Thr Gly Glu Pro
705                 2710                2715                2720

Val Ala Ser Ile Gly Thr Leu Val Leu Arg Ala Val Ser Ala Asp Gln
                2725                2730                2735

Leu Ala Gly Gly Ala Asp Pro Ala Val Arg Asp Ala Leu Phe Arg Val
            2740                2745                2750

Gln Trp Asn Pro Val Arg Leu Pro Pro Ala Gly Ala Ala Val Thr Val
        2755                2760                2765

Ala Thr Leu Gly Ser Leu Ala Gly Ala Pro Phe Asp Gly Tyr Pro Asp
    2770                2775                2780

Leu Ala Ser Leu Ala Arg Ser Gly Arg Val Ala Gly Ala Val Leu Val
785                 2790                2795                2800

Pro Val Glu Ala Gly Ala Gly Glu Val Val Ala Asp Asp Val Val Gly
                2805                2810                2815

Ala Thr His Ala Thr Ala Ala Arg Ala Leu Asp Leu Ala Arg Ser Trp
            2820                2825                2830

Leu Ala Asp Asp Arg Phe Ala Ala Ser Arg Leu Val Phe Val Thr Arg
        2835                2840                2845

Gly Ala Val Ser Gly Ala Asp Leu Ala Gly Ala Ala Val Trp Gly Leu
    2850                2855                2860

Val Arg Ser Ala Leu Ser Glu His Pro Gly Arg Phe Gly Leu Val Asp
865                 2870                2875                2880

Leu Asp Asp Asp Ala Glu Leu Ala Leu Val Pro Arg Val Leu Ala Ser
                2885                2890                2895

Asp Glu Pro Gln Leu Leu Val Arg Gly Gly Glu Val Leu Ala Ala Arg
            2900                2905                2910

Leu Ala Arg Ala Gln Ser Ser His Ala Val Thr Trp Asp Pro Ser Gly
        2915                2920                2925

Thr Val Leu Val Thr Gly Gly Thr Gly Gly Leu Gly Arg Val Met Ala
    2930                2935                2940

Arg His Leu Val Val Glu His Gly Val Arg Asn Leu Leu Leu Val Ser
945                 2950                2955                2960

Arg Arg Gly Pro Ala Ala Glu Gly Ala Glu Glu Leu Val Thr Glu Leu
                2965                2970                2975
```

Arg His Ser Gly Ala Glu Val Ala Val Glu Ala Cys Asp Val Thr Asp
2980 2985 2990

Ala Ala Ala Val Ala Asp Leu Val Ala Arg His Arg Ile Ser Ala Val
2995 3000 3005

Val His Thr Ala Gly Val Leu Asp Asp Gly Val Val Glu Ser Leu Thr
3010 3015 3020

Pro Glu Arg Leu Ser Ala Val Leu Arg Pro Lys Val Asp Ala Ala Trp
025 3030 3035 3040

Asn Leu His Glu Ala Thr Arg Asp Leu Asp Leu Asp Ala Phe Val Val
3045 3050 3055

Phe Ser Ser Val Ala Gly Thr Ile Gly Ser Pro Gly Gln Ala Asn Tyr
3060 3065 3070

Ala Ala Gly Asn Ala Phe Leu Asp Ala Leu Ala His His Arg Arg Ala
3075 3080 3085

Ala Gly Leu Pro Ala Ala Ser Leu Ala Trp Gly Pro Trp Ser Arg Asp
3090 3095 3100

Gly Gly Met Thr Gly Thr Leu Thr Asp Val Asp Ser Ser Ala Ser Pro
105 3110 3115 3120

Gly Arg His Ala Arg Thr His Pro Arg Thr Gly Arg Gly Leu Phe Asp
3125 3130 3135

Ala Ala Leu Ala Ala Gly Asp Ala His Leu Leu Pro Val Arg Phe Asp
3140 3145 3150

Trp Ala Ser Leu Arg Ala Gln Gly Glu Val Pro Pro Leu Leu Arg Gly
3155 3160 3165

Leu Ile Arg Thr Arg Ala Arg Arg Ser Ala Val Gly Gly Ser Ala Ala
3170 3175 3180

Ala Ala Gly Leu Val Gly Arg Leu Ser Gly Arg Gly Thr Val Glu Arg
185 3190 3195 3200

Arg Glu Val Leu Leu Asp Leu Val Arg Ala Gln Ile Ala Val Val Leu
3205 3210 3215

Gly His Ala Asn Pro Glu Thr Ile Glu Ser Thr Arg Val Phe Gln Asp
3220 3225 3230

Leu Gly Phe Asp Ser Leu Thr Ala Val Glu Leu Arg Asn Arg Leu Asn
3235 3240 3245

Asn Ala Thr Gly Leu Arg Leu Ser Ala Thr Ala Val Phe Asp Tyr Pro
3250 3255 3260

Thr Ala Asp Ala Leu Val Asp Phe Leu Leu Asp Glu Leu Phe Gly Ala
265 3270 3275 3280

Gln Glu Glu Ala Glu Leu Pro Ala Pro Val Pro Ser Pro Ala Gly Ala
3285 3290 3295

Ala Asp Asp Pro Val Val Ile Val Gly Met Ser Cys Arg Tyr Pro Gly
3300 3305 3310

Gly Val Gly Ser Pro Glu Asp Leu Trp Arg Leu Val Ser Glu Gly Val

3315 3320 3325

Asp Ala Val Ser Asp Phe Pro Thr Asp Arg Gly Trp Asp Val Glu Ser
3330 3335 3340

Leu Tyr Ser Pro Asp Pro Glu Ala Leu Gly Thr Ser Tyr Thr Arg Ser
345 3350 3355 3360

Gly Gly Phe Leu His Glu Ala Ala Glu Phe Asp Pro Asp Phe Phe Gly
3365 3370 3375

Met Ser Pro Arg Glu Ala Leu Ala Thr Asp Ala Gln Gln Arg Leu Leu
3380 3385 3390

Leu Glu Thr Thr Trp Glu Ala Ile Glu Arg Thr Gly Ile Asp Pro Ala
3395 3400 3405

Ser Leu Arg Gly Ser Arg Thr Gly Val Phe Ala Gly Val Met Tyr Thr
3410 3415 3420

Asp Tyr Gly Asp Leu Leu Val Gly Asp Gln Phe Glu Gly Tyr Arg Ser
425 3430 3435 3440

Asn Gly Ser Ala Ala Ser Ile Ala Ser Gly Arg Val Ser Tyr Thr Phe
3445 3450 3455

Gly Phe Glu Gly Pro Ala Val Thr Val Asp Thr Ala Cys Ser Ser Ser
3460 3465 3470

Leu Val Ala Leu His Trp Ala Ala Gln Ser Leu Arg Ser Gly Glu Cys
3475 3480 3485

Ser Leu Ala Val Ala Gly Gly Val Thr Val Met Ser Thr Pro Thr Thr
3490 3495 3500

Phe Val Glu Phe Ser Arg Gln Arg Gly Leu Ser Ala Asp Gly Arg Cys
505 3510 3515 3520

Lys Ala Phe Ala Asp Ala Ala Asp Gly Val Gly Trp Gly Glu Gly Val
3525 3530 3535

Gly Met Leu Val Leu Glu Arg Leu Ser Asp Ala Arg Arg Asn Gly His
3540 3545 3550

Arg Val Leu Ala Val Val Arg Gly Ser Ala Val Asn Gln Asp Gly Ala
3555 3560 3565

Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ala Gln Gln Arg Val Ile
3570 3575 3580

Arg Gln Ala Leu Ala Ser Ala Gly Leu Ser Ala Ala Asp Val Asp Ala
585 3590 3595 3600

Val Glu Ala His Gly Thr Gly Thr Thr Leu Gly Asp Pro Ile Glu Ala
3605 3610 3615

Gln Ala Leu Leu Ala Thr Tyr Gly Gln Glu Arg Pro Glu Asp Arg Pro
3620 3625 3630

Leu Leu Leu Gly Ser Val Lys Ser Asn Ile Gly His Ala Gln Ala Ala
3635 3640 3645

Ser Gly Val Ala Gly Val Ile Lys Met Val Leu Ala Met Arg His Gly
3650 3655 3660

Val Leu Pro Arg Thr Leu His Val Asp Glu Pro Ser Ser His Val Asp
665 3670 3675 3680

Trp Ser Ala Gly Ala Val Glu Leu Leu Thr Ser Glu Ala Glu Trp Pro
3685 3690 3695

Gln Gly Glu Gly Pro Arg Arg Ala Gly Val Ser Ser Phe Gly Val Ser
3700 3705 3710

Gly Thr Asn Ala His Val Ile Leu Glu Gln Pro Gly Pro Asp Ala Ala
3715 3720 3725

Asp Ala Ala Pro Asp Ala Thr Val Thr Asp Pro Gly Ala Leu Ala Trp
3730 3735 3740

Val Leu Ser Ala Arg Asn Glu Ala Ala Leu Arg Cys Gln Ala Ala Arg
745 3750 3755 3760

Leu Leu Ser Leu Val Ala Gly Ser Asp Ala Leu Cys Ala Arg Asp Ile
3765 3770 3775

Gly His Ser Leu Val Thr Gly Arg Ser Ser Phe Ala His Arg Ala Val
3780 3785 3790

Val Trp Gly Gln Asp Arg Asp Ala Leu Val Arg Ala Leu Ser Ala Leu
3795 3800 3805

Ala Val Gly Glu Ala Asp Ala Gly Leu Ala Glu Gly Ala Ser Gly Ala
3810 3815 3820

Gly Arg Thr Ala Phe Leu Phe Ser Gly Gln Gly Ser Gln Arg Leu Gly
825 3830 3835 3840

Met Gly Trp Glu Leu Tyr Ala Arg Tyr Pro Val Phe Ala Asp Ala Phe
3845 3850 3855

Asp Ala Val Cys Ala Ala Leu Asp Glu His Leu Glu Arg Pro Leu Arg
3860 3865 3870

Asp Val Val Trp Gly Glu Asp Ala Glu Leu Leu Asn Gln Thr Ala Tyr
3875 3880 3885

Ala Gln Ala Gly Leu Phe Ala Ile Glu Val Ala Leu Tyr Arg Leu Ala
3890 3895 3900

Glu Ser Trp Gly Met Arg Pro Asp Phe Val Ala Gly His Ser Ile Gly
905 3910 3915 3920

Glu Val Ala Ala Ala His Val Ser Gly Val Phe Ser Leu Pro Asp Ala
3925 3930 3935

Cys Ala Leu Val Ala Ala Arg Gly Arg Leu Met Gln Gln Leu Pro Ser
3940 3945 3950

Gly Gly Ala Met Met Ala Ile Arg Ala Thr Glu Asp Glu Val Leu Pro
3955 3960 3965

His Leu Ala Glu Gly Val Ser Leu Ala Ala Val Asn Gly Pro Ser Ser
3970 3975 3980

Val Val Ile Ser Gly Ala Glu Asp Ala Val Leu Ala Ile Ala Ala His
985 3990 3995 4000

Phe Ala Gly Glu Gly Arg Lys Thr Thr Arg Leu Arg Val Ser His Ala
4005 4010 4015

Phe His Ser Pro Leu Met Glu Pro Met Leu Glu Glu Phe Arg Ala Val
4020 4025 4030

Val Thr Arg Leu Ser Phe Gly Thr Pro Thr Ile Pro Val Val Ser Asn
4035 4040 4045

Leu Thr Gly Arg Leu Ala Glu Pro Glu Gln Leu Ala His Ala Asp Tyr
4050 4055 4060

Trp Val Arg His Val Arg Glu Ala Val Arg Phe Ala Asp Gly Ile Gln
065 4070 4075 4080

Ala Leu Arg Ala Glu Gly Val Thr Arg Phe Leu Glu Leu Gly Pro Asp
4085 4090 4095

Gly Val Leu Ser Ala Met Ala Arg Glu Ser Ala Ser Asp Asp Ala Val
4100 4105 4110

Leu Ala Pro Val Leu Arg Arg Asp Arg Pro Glu Glu Thr Ala Leu Leu
4115 4120 4125

Gly Ala Leu Ala Gln Leu Tyr Val Arg Gly Ala His Val Asp Trp Thr
4130 4135 4140

Val Pro Phe Ala Gly Ser Gly Ala Arg Trp Ala Asp Leu Pro Thr Tyr
145 4150 4155 4160

Ala Phe Gln His Glu Arg Phe Trp Pro Ser Gly Gly Val Ala Arg Pro
4165 4170 4175

Gly Asp Val Arg Ser Ala Gly Leu Gly Ser Ala Gly His Pro Leu Leu
4180 4185 4190

Gly Ala Ala Val Glu Leu Ala Gly Ser Gly Gly Leu Leu Phe Thr Gly
4195 4200 4205

Arg Leu Ser Val Ser Ser His Pro Trp Leu Ala Asp His Val Val Leu
4210 4215 4220

Gly Ser Val Leu Val Pro Gly Thr Ala Leu Val Glu Leu Val Leu Arg
225 4230 4235 4240

Ala Ala Asp Glu Ala Gly Cys Asp Leu Leu Glu Glu Leu Thr Leu Ala
4245 4250 4255

Ala Pro Leu Val Leu Pro Ala Ser Gly Ala Ala Val Gln Val Gln Val
4260 4265 4270

Ala Val Gly Glu Pro Asp Glu Ala Gly Arg Arg Pro Val Ser Val His
4275 4280 4285

Ala Arg Glu Gly Glu Gly Pro Trp Thr Leu His Ala Ser Gly Ala Val
4290 4295 4300

Thr Ser Gly Ala Glu Val Pro Pro Phe Asp Ala Thr Val Trp Pro Pro
305 4310 4315 4320

Lys Gly Ala Glu Pro Val Asp Val Ala Asp Cys Tyr Asp Val Leu Ala
4325 4330 4335

Asp Ala Gly Leu Thr Tyr Gly Pro Ala Phe His Gly Leu Gln Ala Ala
4340 4345 4350

Trp Lys Leu Gly Gly Asp Val Tyr Ala Glu Ala Lys Leu Pro Glu Ser

4355 4360 4365

Thr Asp Gly Asp Ala Tyr Gly Leu His Pro Ala Leu Phe Asp Ala Ala
4370 4375 4380

Leu His Ala Ser Ala Leu Gly Gly Ala Glu Ala Gly Gly Val Pro Phe
385 4390 4395 4400

Ser Trp Ala Gly Val Ser Leu His Ala Thr Gly Ala Ser His Leu Arg
4405 4410 4415

Val Arg Ile Arg Glu Ala Gly Gly Ala Leu Ser Val Ala Ile Ala Asp
4420 4425 4430

Thr Ser Gly Ala Pro Val Ala Ser Val Glu Ser Leu Val Ile Arg Pro
4435 4440 4445

Leu Ser Ala Gly Gln Val Gln Ala Ala Asp Arg Asp Ala Leu Phe Lys
4450 4455 4460

Ala Asp Trp Val Pro Val Pro Leu Thr Asp Glu Arg Val Glu Pro Gly
465 4470 4475 4480

Thr Gly Pro Glu Gly Glu Pro Leu Arg Thr Tyr Ala Asp Leu Asp Ser
4485 4490 4495

Leu Glu Gly Ala Ala Val Pro Gly Thr Val Leu Val Ala Pro Pro Ser
4500 4505 4510

Gly Ala Ala Gly Thr Val Glu Ser Val His Ala Ala Thr Val Trp Ala
4515 4520 4525

Leu Glu Met Val Gln Ala Trp Leu Ala Asp Asp Arg Phe Ala Thr Ser
4530 4535 4540

Arg Leu Val Phe Val Thr Arg Gly Ala Ala Phe Gly Ala Asp Leu Ala
545 4550 4555 4560

Ala Ala Ala Val Arg Gly Leu Val Arg Ser Ala Gln Ser Glu Asn Pro
4565 4570 4575

Gly Arg Phe Gly Leu Val Asp Met Asp Gly Asp Ala Asp Thr Thr Val
4580 4585 4590

Pro Ala Gln Ala Leu Ala Thr Asp Glu Pro Glu Leu Leu Val Arg Gly
4595 4600 4605

Gly Glu Val Leu Ala Ala Arg Leu Val Arg Ala Gln Ser Ser His Thr
4610 4615 4620

Val Thr Trp Asp Pro Ser Gly Thr Val Leu Ile Thr Gly Gly Thr Gly
625 4630 4635 4640

Gly Leu Gly Arg Ser Val Ala Arg His Leu Val Ser Glu His Gly Val
4645 4650 4655

Arg Ser Leu Leu Leu Val Ser Arg Arg Gly Pro Ala Ala Glu Gly Ala
4660 4665 4670

Gly Glu Leu Val Ala Glu Leu Arg Gly Ser Gly Ala Glu Val Val Ile
4675 4680 4685

Glu Ala Cys Asp Val Thr Asp Ala Val Ala Val Ala Asp Leu Val Ala
4690 4695 4700

Arg His Arg Ile Ser Ala Val Val His Thr Ala Gly Val Leu Asp Asp
705 4710 4715 4720

Gly Val Val Glu Ser Leu Thr Pro Glu Arg Leu Ala Val Val Leu Arg
4725 4730 4735

Pro Lys Val Asp Ala Ala Trp Asn Leu His Glu Ala Thr Arg Gly Leu
4740 4745 4750

Asp Leu Asp Ala Phe Val Val Phe Ser Ser Val Ala Gly Thr Phe Gly
4755 4760 4765

Ser Ala Gly Gln Ala Asn Tyr Ala Ala Gly Asn Ala Phe Leu Asp Ala
4770 4775 4780

Leu Ala Tyr His Arg Arg Ala Val Gly Leu Pro Ala Val Ser Leu Ala
785 4790 4795 4800

Trp Gly Pro Trp Ser Gln Asp Gly Gly Met Thr Gly Thr Leu Ser Asp
4805 4810 4815

Ala Asp Val Gln Arg Ile Ala Arg Gln Gly Met Pro Pro Leu Thr Val
4820 4825 4830

Glu Glu Gly Leu Ala Leu Phe Asp Ala Ala Leu Gly Ser Ala Glu Pro
4835 4840 4845

Met Ala Leu Pro Val Arg Leu Asp Leu Ala Ala Leu Arg Ala Gln Gly
4850 4855 4860

Glu Pro Gln Pro Leu Leu Arg Gly Leu Ile Arg Thr Arg Thr Arg Arg
865 4870 4875 4880

Ser Gly Ala Ala Ala Ala Ser Gly Ile Ala Gln Arg Leu Ala Gly Leu
4885 4890 4895

Ser Thr Ala Glu Arg Arg Glu Ala Leu Leu Asp Val Val Arg Ala Gln
4900 4905 4910

Ile Ala Thr Val Leu Gly His Ala Gly Pro Glu Thr Ile Ala Pro Asp
4915 4920 4925

Arg Ala Phe Gln Asp Leu Gly Leu Asp Ser Leu Thr Ala Ile Glu Leu
4930 4935 4940

Arg Asn Leu Leu Gly Lys Ala Thr Gly Leu Arg Leu Pro Ala Thr Thr
945 4950 4955 4960

Val Phe Asp Tyr Pro Thr Val Asp Ala Leu Ala Ala His Leu Leu Asp
4965 4970 4975

Glu Leu Phe Gly Ala Glu Thr Gly Thr Ala Thr Glu Thr Pro Leu Pro
4980 4985 4990

Val Pro Gly Leu Pro Ser Leu Ala Asp Asp Pro Val Val Ile Val Gly
4995 5000 5005

Met Ser Cys Arg Phe Pro Gly Gly Val Ala Ser Pro Glu Asp Leu Trp
5010 5015 5020

Arg Leu Val Ala Asp Gly Val Asp Ala Val Ser Ala Phe Pro Thr Asp
025 5030 5035 5040

Arg Gly Trp Glu Ile Asp Asp Thr Tyr Asp Pro Glu Arg Glu Gly Ala
5045 5050 5055

Ile Ala Thr Arg Ser Gly Gly Phe Leu His Asp Ala Ala Glu Phe Asp
5060 5065 5070

Pro Glu Phe Phe Gly Met Ser Pro Arg Glu Ala Leu Thr Thr Asp Ala
5075 5080 5085

Gln Gln Arg Leu Leu Leu Glu Thr Thr Trp Glu Ala Leu Glu Arg Ala
5090 5095 5100

Gly Met Asp Pro Ala Thr Leu Arg Gly Ser Arg Thr Gly Val Phe Ala
105 5110 5115 5120

Gly Val Met Tyr His Asp Tyr Ser Thr Leu Leu Ser Gly Arg Glu Phe
5125 5130 5135

Glu Gly Tyr Gln Gly Ser Gly Ser Ala Gly Ser Val Ala Ser Gly Arg
5140 5145 5150

Val Ser Tyr Thr Phe Gly Phe Glu Gly Pro Ala Val Thr Val Asp Thr
5155 5160 5165

Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala Ala Gln Ser Leu
5170 5175 5180

Arg Ser Gly Glu Cys Ser Leu Ala Leu Ala Gly Gly Val Thr Val Met
185 5190 5195 5200

Ser Thr Pro Leu Thr Phe Val Glu Phe Ser Arg Gln Gly Gly Leu Ser
5205 5210 5215

Ala Asp Gly Arg Cys Lys Ala Phe Ala Asp Ala Ala Asp Gly Val Gly
5220 5225 5230

Trp Ala Glu Gly Ala Gly Ile Leu Val Leu Glu Arg Leu Ser Asp Ala
5235 5240 5245

Arg Arg Asn Gly His Arg Ile Leu Ala Thr Val Arg Gly Ser Ala Val
5250 5255 5260

Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ala
265 5270 5275 5280

Gln Gln Arg Val Ile Arg Gln Ala Leu Ala Ser Ala Gly Leu Ser Ala
5285 5290 5295

Ala Asp Val Asp Ala Val Glu Ala His Gly Thr Gly Thr Thr Leu Gly
5300 5305 5310

Asp Pro Ile Glu Ala Gln Ala Leu Leu Ala Thr Tyr Gly Gln Glu Arg
5315 5320 5325

Pro Glu Asp Arg Pro Leu Leu Leu Gly Ser Val Lys Ser Asn Ile Gly
5330 5335 5340

His Ala Gln Ala Ala Ser Gly Val Ala Gly Val Ile Lys Met Val Leu
345 5350 5355 5360

Ala Met Arg His Gly Val Leu Pro Arg Thr Leu His Val Asp Glu Pro
5365 5370 5375

Ser Ser His Val Asp Trp Ser Ala Gly Ala Val Glu Leu Leu Thr Ser
5380 5385 5390

Glu Ala Glu Trp Pro Gln Gly Glu Gly Pro Arg Arg Ala Gly Val Ser

5395 5400 5405

Ser Phe Gly Ile Ser Gly Thr Asn Ala His Val Ile Leu Glu Gln Pro
5410 5415 5420

Glu Pro Val Ala Ala Glu Thr Glu Ser Ile Thr Pro Asp Thr Ala Pro
425 5430 5435 5440

Asp Ala Ala Glu Asp Glu Ala Ala Asp Ser Gly Thr Pro Val Pro Ala
5445 5450 5455

Leu Leu Ser Gly Arg Ser Ala Ser Ala Leu Arg Ala Gln Ala Ala Arg
5460 5465 5470

Leu Leu Ser Arg Leu Asp Gly Asp Pro Gly Pro Arg Ile Thr Asp Val
5475 5480 5485

Ala Tyr Ser Leu Ala Thr Gly Arg Ser Ala Phe Pro His Arg Ala Val
5490 5495 5500

Ile Leu Ala Ala Asn Arg Ala Asp Leu Leu His Ser Leu Ser Ala Leu
505 5510 5515 5520

Ala Glu Gly His Thr Glu Ala Pro Ala Val Val Ala Gln Asp Arg Ala
5525 5530 5535

Arg Ser Gly Lys Leu Ala Phe Leu Phe Ser Gly Gln Gly Ser Gln Arg
5540 5545 5550

Leu Gly Met Gly Arg Glu Leu Tyr Gly Arg Tyr Pro Ala Phe Ala Glu
5555 5560 5565

Ala Leu Asp Ala Val Cys Ala Ala Leu Asp Ala His Leu Asp Arg Pro
5570 5575 5580

Leu Arg Asp Val Ile Trp Gly Glu Asp Ala Glu Leu Leu Asn Arg Thr
585 5590 5595 5600

Gly Tyr Ala Gln Thr Gly Leu Phe Ala Ile Glu Val Ala Leu Phe Arg
5605 5610 5615

Leu Leu Glu Ser Trp Gly Val Arg Pro Asp His Leu Leu Gly His Ser
5620 5625 5630

Ile Gly Glu Ile Ala Ala Ala His Val Ala Gly Val Leu Ser Leu Pro
5635 5640 5645

Asp Ala Cys Ala Leu Val Ala Ala Arg Gly Arg Leu Met Gln Gln Leu
5650 5655 5660

Pro Ser Gly Gly Ala Met Met Ala Ile Arg Ala Thr Glu Asp Glu Val
665 5670 5675 5680

Leu Pro His Leu Ala Glu Gly Val Ser Leu Ala Ala Val Asn Gly Pro
5685 5690 5695

Ser Ser Val Val Val Ser Gly Ala Glu Asp Glu Val Leu Ala Leu Ala
5700 5705 5710

Ala His Phe Glu Glu Glu Gly Arg Lys Thr Thr Arg Leu Arg Val Ser
5715 5720 5725

His Ala Phe His Ser Pro Leu Met Glu Pro Met Leu Ala Asp Phe Arg
5730 5735 5740

Ala Val Ala Asp Gly Met Thr Tyr Ala Ala Pro Arg Ile Pro Val Val
745 5750 5755 5760

Ser Asn Val Thr Gly Arg Pro Ala Thr Ala Glu Glu Leu Cys Cys Ala
5765 5770 5775

Glu Tyr Trp Val Gly His Val Arg Glu Ala Val Arg Phe Ala Asp Gly
5780 5785 5790

Val Gly Ala Leu Arg Glu Gln Gly Val Thr Thr Phe Leu Glu Leu Gly
5795 5800 5805

Pro Asp Gly Ser Leu Ser Ala Leu Ala Ala Glu Ser Ala Ala Asp Asp
5810 5815 5820

Ser Val Leu Ala Pro Val Leu Arg Lys Asn Arg Pro Glu Ala Pro Ala
825 5830 5835 5840

Leu Leu Thr Ala Leu Ala Arg Leu His Ala Gln Gly Thr Pro Val Asp
5845 5850 5855

Trp Ser Ala Ala Phe Ala Gly Thr Gly Ala Arg Trp Val Asp Leu Pro
5860 5865 5870

Thr Tyr Ala Phe Gln His Glu Arg Phe Trp Pro Ser Gly Gly Ala Ala
5875 5880 5885

Arg Ala Gly Asp Val Arg Ser Ala Gly Leu Gly Ser Ala Gly His Pro
5890 5895 5900

Leu Leu Gly Ala Ala Val Glu Leu Ala Gly Ser Gly Gly Arg Leu Leu
905 5910 5915 5920

Thr Gly Arg Leu Ser Leu Ser Ser His Pro Trp Leu Ala Asp His Val
5925 5930 5935

Val Leu Gly Ser Val Leu Val Pro Gly Thr Ala Leu Met Glu Leu Val
5940 5945 5950

Leu Arg Ala Ala Asp Glu Val Asp Cys Ala Ala Val Asp Glu Leu Thr
5955 5960 5965

Leu Ala Ala Pro Leu Val Leu Pro Ala Ser Gly Ala Ala Ile Gln Val
5970 5975 5980

Gln Val Trp Val Gly Glu Pro Asp Glu Ala Gly Arg Arg Pro Val Ser
985 5990 5995 6000

Val His Ala Arg Glu Gly Glu Gly Pro Trp Thr Leu His Ala Asp Gly
6005 6010 6015

Ala Leu Ala Pro Ala Ala Glu Thr Val Pro Phe Asp Thr Ala Ile Trp
6020 6025 6030

Pro Pro Gln Gly Ala Glu His Leu Asp Ala Ala Gly Cys Tyr Glu Arg
6035 6040 6045

Phe Ala Asp Ala Gly Phe Ala Tyr Gly Pro Val Phe Gln Gly Leu Arg
6050 6055 6060

Ala Ala Trp Lys Leu Gly Glu Asp Ile Tyr Ala Glu Val Ala Leu Pro
065 6070 6075 6080

Glu Gly Thr Asp Gly Asn Ala Tyr Gly Leu His Pro Ala Leu Phe Asp
6085 6090 6095

Ala Ala Leu His Ala Ala Leu Leu Gly Gly Glu Gly Thr Asp Glu Ala
6100 6105 6110

Ala Val Pro Phe Ser Trp Asn Gly Val Thr Leu His Ala Thr Gly Ala
6115 6120 6125

Ser Arg Val Arg Val Arg Ile Arg Pro Thr Glu Gly Gly Thr Ser Ile
6130 6135 6140

Ala Leu Val Asp Thr Ala Gly Ala Pro Val Ala Ser Val Arg Ser Leu
145 6150 6155 6160

Thr Ala Arg Pro Ile Thr Ala Gly Gln Leu Gln Thr Gly Asp Arg Asp
6165 6170 6175

Ser Leu Phe Gln Val Asp Trp Thr Thr Leu His Leu Thr Asp Glu Arg
6180 6185 6190

Ala Asn Ser Leu Ala Leu Leu Gly Lys Asp Thr Glu Gly Ile Leu Asp
6195 6200 6205

Thr Leu Ser Leu Gln Pro His Ala Asp Leu Asp Asp Leu Ala Ala Thr
6210 6215 6220

Gly Val His Asp Thr Val Leu Ala Pro Leu Pro Thr Arg Thr Ala Gly
225 6230 6235 6240

Thr Val Glu Ser Val His Ala Ala Thr Thr Gly Ala Leu Ala Leu Ile
6245 6250 6255

Arg Ser Trp Leu Ala Asp Asp Arg Phe Ala Ala Ser Arg Leu Val Phe
6260 6265 6270

Val Thr Arg Gly Ala Val Ser Gly Thr Asp Leu Ala Gly Ala Ser Val
6275 6280 6285

Trp Gly Leu Val Arg Ser Ala Leu Leu Glu His Pro Gly Arg Phe Gly
6290 6295 6300

Leu Val Asp Val Asp Val Asp Gln Asp Ala Glu Val Pro Leu Val Pro
305 6310 6315 6320

Arg Ala Leu Ala Ser Asp Glu Pro Gln Val Leu Val Arg Gly Gly Glu
6325 6330 6335

Val Leu Ala Ala Arg Leu Val Arg Ala Gln Ser Ser Asp Thr Val Thr
6340 6345 6350

Trp Asp Pro Ser Gly Thr Val Leu Ile Thr Gly Gly Thr Gly Gly Leu
6355 6360 6365

Gly Arg Ser Val Ala Arg His Leu Val Ser Glu His Gly Val Arg Ser
6370 6375 6380

Leu Leu Leu Val Ser Arg Arg Gly Pro Ala Ala Glu Gly Val Asp Ala
385 6390 6395 6400

Leu Val Ala Glu Leu Ala Glu Cys Gly Ala Gln Val Thr Val Glu Ala
6405 6410 6415

Cys Asp Val Thr Asp Ala Val Ala Val Ala Asp Leu Val Ala Arg His
6420 6425 6430

Arg Ile Ser Ala Val Val His Thr Ala Gly Val Leu Asp Asp Gly Val

6435 6440 6445

Val Glu Ser Leu Thr Pro Glu Arg Leu Ser Ala Val Leu Arg Pro Lys
6450 6455 6460

Val Asp Ala Ala Trp Asn Leu His Glu Ala Thr Arg Gly Leu Asp Leu
465 6470 6475 6480

Asp Ala Phe Val Val Phe Ser Ser Val Ala Gly Thr Phe Gly Ser Ala
6485 6490 6495

Gly Gln Ala Asn Tyr Ala Ala Gly Asn Ala Phe Leu Asp Ala Leu Ala
6500 6505 6510

Tyr His Arg Arg Ala Val Gly Leu Pro Ala Val Ser Leu Ala Trp Gly
6515 6520 6525

Pro Trp Ser Gln Asp Gly Gly Met Thr Gly Thr Leu Ser Asp Ala Asp
6530 6535 6540

Val Gln Arg Ile Ala Arg Gln Gly Met Pro Pro Leu Thr Val Glu Glu
545 6550 6555 6560

Gly Leu Ala Leu Phe Asp Ala Ala Leu Gly Ser Ala Glu Pro Met Ala
6565 6570 6575

Leu Pro Val Arg Leu Asp Leu Ala Ala Leu Arg Ala Gln Gly Glu Pro
6580 6585 6590

Gln Pro Leu Leu Arg Gly Leu Ile Arg Thr Pro Gly Arg Arg Thr Ala
6595 6600 6605

Ala Ala Ala Thr Glu Gly Asp Thr Ala Ala Ala Phe Ala Gly Arg Leu
6610 6615 6620

Thr Gly Leu Ser Ala Ala Glu Gly Arg Glu Val Val Leu Gly Ala Val
625 6630 6635 6640

Arg Ser Gln Ile Ala Gly Val Leu Gly His Ala Glu Ala Thr Glu Ile
6645 6650 6655

Asp Gln Asp Arg Ala Phe Leu Asp Leu Gly Phe Asp Ser Leu Thr Ala
6660 6665 6670

Val Glu Leu Arg Asn Arg Leu Gly Ala Val Thr Gly Ile Arg Leu Pro
6675 6680 6685

Ala Thr Leu Leu Phe Asp Tyr Pro Thr Pro Ala Glu Leu Val Ala His
6690 6695 6700

Leu His Ala Arg Ile Ala Pro Glu Pro Thr Val Gly Pro Glu Ala Leu
705 6710 6715 6720

Leu Gly Glu Leu Glu Arg Met Glu Lys Ser Phe Gly Gly Leu Asp Leu
6725 6730 6735

Thr Glu Glu Met His Glu Gln Ile Ala Gly Arg Leu Glu Val Leu Arg
6740 6745 6750

Ala Lys Trp Asp Ala Leu Arg Asp Thr Ala Ala Ala Ala Gly His Asp
6755 6760 6765

Gly Ser Pro Ser Asp Glu Asp Phe Asp Phe Glu Ser Ala Ser Asp Asp
6770 6775 6780

```
Glu Val Phe Asp Leu Leu Asp Asn Glu Leu Gly Leu Ser
785                 6790                 6795


<210> 5
<211> 1650
<212> DNA
<213> Streptomyces natalensis

<220>
<221> CDS
<222> (1)..(1650)
<223> ORF1

<400> 5
atg ttc gag aac cag cat ctg tca cgg cgc cgt ctg ctc gga ctg gcc   48
Met Phe Glu Asn Gln His Leu Ser Arg Arg Arg Leu Leu Gly Leu Ala
  1               5                  10                  15

gcc ctc ggc ggc gcc gca gcc gcc gga atg acc acg atc acc tcc gcc   96
Ala Leu Gly Gly Ala Ala Ala Ala Gly Met Thr Thr Ile Thr Ser Ala
            20                  25                  30

cct cat gcc gcg gcc gcc gac cgg cgc agt ccg cag gcc cgc agc ggc   144
Pro His Ala Ala Ala Ala Asp Arg Arg Ser Pro Gln Ala Arg Ser Gly
        35                  40                  45

tcg ttc gta ccg gcc gtg gtg atc ggt acg gga tac ggc gcc gcg gtc   192
Ser Phe Val Pro Ala Val Val Ile Gly Thr Gly Tyr Gly Ala Ala Val
    50                  55                  60

tcc gcg ctg cgg ctc ggc gag gcc gga att ccc acg ctc atg ctc gaa   240
Ser Ala Leu Arg Leu Gly Glu Ala Gly Ile Pro Thr Leu Met Leu Glu
65                  70                  75                  80

atg ggc cag ctg tgg aac aag ccc gcc gac gac ggc aac gtc ttc tgc   288
Met Gly Gln Leu Trp Asn Lys Pro Ala Asp Asp Gly Asn Val Phe Cys
                85                  90                  95

gga atg ctc tcg ccc gac cgc cgc tcc agc tgg ttc aag tcc cgc acc   336
Gly Met Leu Ser Pro Asp Arg Arg Ser Ser Trp Phe Lys Ser Arg Thr
            100                 105                 110

gag gcc ccg ctc ggc tcg ttc ctg tgg ctg gat gtg atc aac cgc gac   384
Glu Ala Pro Leu Gly Ser Phe Leu Trp Leu Asp Val Ile Asn Arg Asp
        115                 120                 125

atc gac ccg tac gcg gga gtg ctg gac aag gtg cac ttc gac cag atg   432
Ile Asp Pro Tyr Ala Gly Val Leu Asp Lys Val His Phe Asp Gln Met
    130                 135                 140

tcg gtg tac gtg ggg cgg ggt gtc ggc ggc ggc tcg ctg gtc aac ggc   480
Ser Val Tyr Val Gly Arg Gly Val Gly Gly Gly Ser Leu Val Asn Gly
145                 150                 155                 160

ggg atg gcc gtc gta ccg aag cgc tcg tac ttc gag gag gtc ctc ccg   528
Gly Met Ala Val Val Pro Lys Arg Ser Tyr Phe Glu Glu Val Leu Pro
                165                 170                 175

cgg gtg gac gcc gcc gag atg tac gac cgg tac ttc ccg cgc gcc aac   576
Arg Val Asp Ala Ala Glu Met Tyr Asp Arg Tyr Phe Pro Arg Ala Asn
            180                 185                 190

tcc atg ctc aag gtg aac cac atc gac aag ggg tgg ttc gag gag acg   624
Ser Met Leu Lys Val Asn His Ile Asp Lys Gly Trp Phe Glu Glu Thr
```

```
        195                 200                 205

gag tgg tac aag ttc gcg cgg gtc tcg cgc gag cag gcg ggc aag gcg   672
Glu Trp Tyr Lys Phe Ala Arg Val Ser Arg Glu Gln Ala Gly Lys Ala
    210                 215                 220

ggc ctg ggc acc acc ttc gtc ccc aac gtc tac gac ttc gac tac atg   720
Gly Leu Gly Thr Thr Phe Val Pro Asn Val Tyr Asp Phe Asp Tyr Met
225                 230                 235                 240

cgg cgc gag gcg aac ggt gag tcg ccc aag tcc gcg ctg gcg acc gag   768
Arg Arg Glu Ala Asn Gly Glu Ser Pro Lys Ser Ala Leu Ala Thr Glu
                245                 250                 255

gtc atc tac ggc aac aac cac ggc aaa cag agc ctg gac aag acc tac   816
Val Ile Tyr Gly Asn Asn His Gly Lys Gln Ser Leu Asp Lys Thr Tyr
            260                 265                 270

ctg gcc gcc gcg ctc ggc acc ggc aag gtc acc atc gag acc ctg cac   864
Leu Ala Ala Ala Leu Gly Thr Gly Lys Val Thr Ile Glu Thr Leu His
        275                 280                 285

cag gtc agg gcg atc cac cag cag ccg gac ggc agc tac gtg ctg tcc   912
Gln Val Arg Ala Ile His Gln Gln Pro Asp Gly Ser Tyr Val Leu Ser
    290                 295                 300

gtg gac cag atc gac acg gcc ggc cag acc gtc gcc cac aag gag atc   960
Val Asp Gln Ile Asp Thr Ala Gly Gln Thr Val Ala His Lys Glu Ile
305                 310                 315                 320

tcc tgc cgt cac ctg ttc ctc ggc gcc ggc agc ctc ggc tcc acc gaa  1008
Ser Cys Arg His Leu Phe Leu Gly Ala Gly Ser Leu Gly Ser Thr Glu
                325                 330                 335

ctg ctg gtg cgc gcc cgg gac acc ggc gcg ctg ccc gac ctc aac gcc  1056
Leu Leu Val Arg Ala Arg Asp Thr Gly Ala Leu Pro Asp Leu Asn Ala
            340                 345                 350

gag gtc ggc gcg ggc tgg ggc ccc aac ggc aac atc atg acc ggc cgg  1104
Glu Val Gly Ala Gly Trp Gly Pro Asn Gly Asn Ile Met Thr Gly Arg
        355                 360                 365

gcc aac cac gtc tgg aac ccc acc ggg gcc cac cag tcc tcg atc ccc  1152
Ala Asn His Val Trp Asn Pro Thr Gly Ala His Gln Ser Ser Ile Pro
    370                 375                 380

gct ctg ggc atc gac gac tgg aac aac ccc acc gcc ccg gtc ttc gcc  1200
Ala Leu Gly Ile Asp Asp Trp Asn Asn Pro Thr Ala Pro Val Phe Ala
385                 390                 395                 400

gaa atc gcc ccg atg ccc gcc ggg ttg gag acc tgg gtc agc ctc tat  1248
Glu Ile Ala Pro Met Pro Ala Gly Leu Glu Thr Trp Val Ser Leu Tyr
                405                 410                 415

ctg gcg atc acc aag aac ccc gag cgc ggc acc ttc gtc tac gac aag  1296
Leu Ala Ile Thr Lys Asn Pro Glu Arg Gly Thr Phe Val Tyr Asp Lys
            420                 425                 430

gcc acc gac cgg gcc gcg ctg cgc tgg acg cgg gac cag aac acg ccc  1344
Ala Thr Asp Arg Ala Ala Leu Arg Trp Thr Arg Asp Gln Asn Thr Pro
        435                 440                 445

gcg gtc aac gcc gcc agg tcg ctc ttc gac cgc atc aac aag gcc aac  1392
Ala Val Asn Ala Ala Arg Ser Leu Phe Asp Arg Ile Asn Lys Ala Asn
    450                 455                 460
```

```
ggc acg atg tac cgc tac gac ctg ttc ggg ccg cag ctg aag aac ttc   1440
Gly Thr Met Tyr Arg Tyr Asp Leu Phe Gly Pro Gln Leu Lys Asn Phe
465                 470                 475                 480

tcc gac gac ttc tgc tac cac ccg ctc ggc ggc tgc gtc ctg ggc aag   1488
Ser Asp Asp Phe Cys Tyr His Pro Leu Gly Gly Cys Val Leu Gly Lys
                485                 490                 495

gcc acc gac ggg tac ggc cgg gtc gcc ggc tac cac aac ctc tac gtc   1536
Ala Thr Asp Gly Tyr Gly Arg Val Ala Gly Tyr His Asn Leu Tyr Val
            500                 505                 510

acg gac ggc gcg ctc atc ccg ggg tcc atc ggg gtc aac ccc ttc gtg   1584
Thr Asp Gly Ala Leu Ile Pro Gly Ser Ile Gly Val Asn Pro Phe Val
        515                 520                 525

acc atc acg gcg ctg gcc gag cgg aac atc gag cgg atc atc gcg gag   1632
Thr Ile Thr Ala Leu Ala Glu Arg Asn Ile Glu Arg Ile Ile Ala Glu
    530                 535                 540

gac gtc aag gcc gcc tag                                          1650
Asp Val Lys Ala Ala
545                 550
```

<210> 6
<211> 550
<212> PRT
<213> Streptomyces natalensis

<400> 6

```
Met Phe Glu Asn Gln His Leu Ser Arg Arg Arg Leu Leu Gly Leu Ala
  1               5                  10                  15

Ala Leu Gly Gly Ala Ala Ala Ala Gly Met Thr Thr Ile Thr Ser Ala
            20                  25                  30

Pro His Ala Ala Ala Ala Asp Arg Arg Ser Pro Gln Ala Arg Ser Gly
        35                  40                  45

Ser Phe Val Pro Ala Val Val Ile Gly Thr Gly Tyr Gly Ala Ala Val
    50                  55                  60

Ser Ala Leu Arg Leu Gly Glu Ala Gly Ile Pro Thr Leu Met Leu Glu
65                  70                  75                  80

Met Gly Gln Leu Trp Asn Lys Pro Ala Asp Asp Gly Asn Val Phe Cys
                85                  90                  95

Gly Met Leu Ser Pro Asp Arg Arg Ser Ser Trp Phe Lys Ser Arg Thr
            100                 105                 110

Glu Ala Pro Leu Gly Ser Phe Leu Trp Leu Asp Val Ile Asn Arg Asp
        115                 120                 125

Ile Asp Pro Tyr Ala Gly Val Leu Asp Lys Val His Phe Asp Gln Met
    130                 135                 140

Ser Val Tyr Val Gly Arg Gly Val Gly Gly Gly Ser Leu Val Asn Gly
145                 150                 155                 160

Gly Met Ala Val Val Pro Lys Arg Ser Tyr Phe Glu Glu Val Leu Pro
                165                 170                 175
```

Arg Val Asp Ala Ala Glu Met Tyr Asp Arg Tyr Phe Pro Arg Ala Asn
180 185 190

Ser Met Leu Lys Val Asn His Ile Asp Lys Gly Trp Phe Glu Glu Thr
195 200 205

Glu Trp Tyr Lys Phe Ala Arg Val Ser Arg Glu Gln Ala Gly Lys Ala
210 215 220

Gly Leu Gly Thr Thr Phe Val Pro Asn Val Tyr Asp Phe Asp Tyr Met
225 230 235 240

Arg Arg Glu Ala Asn Gly Glu Ser Pro Lys Ser Ala Leu Ala Thr Glu
245 250 255

Val Ile Tyr Gly Asn Asn His Gly Lys Gln Ser Leu Asp Lys Thr Tyr
260 265 270

Leu Ala Ala Ala Leu Gly Thr Gly Lys Val Thr Ile Glu Thr Leu His
275 280 285

Gln Val Arg Ala Ile His Gln Gln Pro Asp Gly Ser Tyr Val Leu Ser
290 295 300

Val Asp Gln Ile Asp Thr Ala Gly Gln Thr Val Ala His Lys Glu Ile
305 310 315 320

Ser Cys Arg His Leu Phe Leu Gly Ala Gly Ser Leu Gly Ser Thr Glu
325 330 335

Leu Leu Val Arg Ala Arg Asp Thr Gly Ala Leu Pro Asp Leu Asn Ala
340 345 350

Glu Val Gly Ala Gly Trp Gly Pro Asn Gly Asn Ile Met Thr Gly Arg
355 360 365

Ala Asn His Val Trp Asn Pro Thr Gly Ala His Gln Ser Ser Ile Pro
370 375 380

Ala Leu Gly Ile Asp Asp Trp Asn Asn Pro Thr Ala Pro Val Phe Ala
385 390 395 400

Glu Ile Ala Pro Met Pro Ala Gly Leu Glu Thr Trp Val Ser Leu Tyr
405 410 415

Leu Ala Ile Thr Lys Asn Pro Glu Arg Gly Thr Phe Val Tyr Asp Lys
420 425 430

Ala Thr Asp Arg Ala Ala Leu Arg Trp Thr Arg Asp Gln Asn Thr Pro
435 440 445

Ala Val Asn Ala Ala Arg Ser Leu Phe Asp Arg Ile Asn Lys Ala Asn
450 455 460

Gly Thr Met Tyr Arg Tyr Asp Leu Phe Gly Pro Gln Leu Lys Asn Phe
465 470 475 480

Ser Asp Asp Phe Cys Tyr His Pro Leu Gly Gly Cys Val Leu Gly Lys
485 490 495

Ala Thr Asp Gly Tyr Gly Arg Val Ala Gly Tyr His Asn Leu Tyr Val
500 505 510

Thr Asp Gly Ala Leu Ile Pro Gly Ser Ile Gly Val Asn Pro Phe Val
515 520 525

```
Thr Ile Thr Ala Leu Ala Glu Arg Asn Ile Glu Arg Ile Ile Ala Glu
    530                 535                 540

Asp Val Lys Ala Ala
545                 550
```

<210> 7
<211> 1197
<212> DNA
<213> Streptomyces natalensis

<220>
<221> CDS
<222> (1)..(1197)
<223> ORF2

<400> 7

```
atg acg tac aca gac ccg gcc gcg ccc gag acg gat ccg ccg gcc gtc   48
Met Thr Tyr Thr Asp Pro Ala Ala Pro Glu Thr Asp Pro Pro Ala Val
  1               5                  10                  15

gac ttt ccg cag cgc aag ccc ggc gtg ccg ttc ccg ccg ccc gac tac   96
Asp Phe Pro Gln Arg Lys Pro Gly Val Pro Phe Pro Pro Pro Asp Tyr
            20                  25                  30

gcc gac tac cgc gac cgg aag ggg ctc gtc ctc tcg cag ctg tcc gac  144
Ala Asp Tyr Arg Asp Arg Lys Gly Leu Val Leu Ser Gln Leu Ser Asp
        35                  40                  45

ggc aaa cgg gta tgg ctg gtc acc cgg cac gag gac gta cgc gcc gta  192
Gly Lys Arg Val Trp Leu Val Thr Arg His Glu Asp Val Arg Ala Val
    50                  55                  60

ctg acc agc ccg agc atc agc tcg aac ccc gag cac aag gga ttt ccc  240
Leu Thr Ser Pro Ser Ile Ser Ser Asn Pro Glu His Lys Gly Phe Pro
65                  70                  75                  80

aac gtc ggg aac ctg ggt gtg ccc aag cag gac cag atc ccg ggc tgg  288
Asn Val Gly Asn Leu Gly Val Pro Lys Gln Asp Gln Ile Pro Gly Trp
                85                  90                  95

ttc gtg ggc atg gac tcc ccc gag cac gac cgg ttc cgc aag gcc ctc  336
Phe Val Gly Met Asp Ser Pro Glu His Asp Arg Phe Arg Lys Ala Leu
            100                 105                 110

atc ccg gag ttc acc gtc cgg cgg gta cgc gcg atg aag ccc gcg atc  384
Ile Pro Glu Phe Thr Val Arg Arg Val Arg Ala Met Lys Pro Ala Ile
        115                 120                 125

gaa cgc acg gtg gac gcc caa ctg gac gcc atg ctg gcc gcg ggc aac  432
Glu Arg Thr Val Asp Ala Gln Leu Asp Ala Met Leu Ala Ala Gly Asn
    130                 135                 140

acc gcc gac ctc gtc gcc gac ttc gcc ctg ccc atc ccc tcc ctg gtg  480
Thr Ala Asp Leu Val Ala Asp Phe Ala Leu Pro Ile Pro Ser Leu Val
145                 150                 155                 160

atc tcc gca ctg ctc ggc gtg ccg ccc gcc gac cgc gag ttc ttc gag  528
Ile Ser Ala Leu Leu Gly Val Pro Pro Ala Asp Arg Glu Phe Phe Glu
                165                 170                 175

tcc agg acc cgc gtc ctg gtc tcc ctc cgc tcc tcc acc gac gac gac  576
Ser Arg Thr Arg Val Leu Val Ser Leu Arg Ser Ser Thr Asp Asp Asp
```

```
                180                 185                 190

cgg atg gcc gcc gcc aag gac ctc ctg cgg tac atc aac cgg ctc gtg   624
Arg Met Ala Ala Ala Lys Asp Leu Leu Arg Tyr Ile Asn Arg Leu Val
        195                 200                 205

gag atc aaa cag aag tgg ggc ggc gac gac ctc atc acc cgg ctg ctg   672
Glu Ile Lys Gln Lys Trp Gly Gly Asp Asp Leu Ile Thr Arg Leu Leu
    210                 215                 220

gcc acc ggt gcc atc gcc ccc cac gaa atg tcc ggc gtg ctg atg ctc   720
Ala Thr Gly Ala Ile Ala Pro His Glu Met Ser Gly Val Leu Met Leu
225                 230                 235                 240

ctg ctc atc gcc ggc cac gag acc acg gcc aac aac atc gcc ctc ggc   768
Leu Leu Ile Ala Gly His Glu Thr Thr Ala Asn Asn Ile Ala Leu Gly
                245                 250                 255

gtg gtc acc ctg ctg gcg aac ccc caa tgg atc ggc gac gac cgg gcc   816
Val Val Thr Leu Leu Ala Asn Pro Gln Trp Ile Gly Asp Asp Arg Ala
            260                 265                 270

gtg gag gag acc ctg cgc ttc cac tcc gtc gcc gac ctg gtg tcc ctg   864
Val Glu Glu Thr Leu Arg Phe His Ser Val Ala Asp Leu Val Ser Leu
        275                 280                 285

cgc gtc gcg gtc cag gac gtg gaa atc gcc ggg cag ctc atc aag gcg   912
Arg Val Ala Val Gln Asp Val Glu Ile Ala Gly Gln Leu Ile Lys Ala
    290                 295                 300

ggc gag gga atc gtg ccg ctg gtc gcc gcc gcc aat cat gac gag aac   960
Gly Glu Gly Ile Val Pro Leu Val Ala Ala Ala Asn His Asp Glu Asn
305                 310                 315                 320

gcc ttc gaa tgc ccc cac gcc ttc gac ccg tcc cgg tcc gcc cgc cac  1008
Ala Phe Glu Cys Pro His Ala Phe Asp Pro Ser Arg Ser Ala Arg His
                325                 330                 335

cat gtg gcc ttc ggc tac ggc gta cac caa tgc ctg gga cag aac ctg  1056
His Val Ala Phe Gly Tyr Gly Val His Gln Cys Leu Gly Gln Asn Leu
            340                 345                 350

gtg cgg atc gag atg gaa gtc gcg tac cgg aaa ctc ttc gag cgc atc  1104
Val Arg Ile Glu Met Glu Val Ala Tyr Arg Lys Leu Phe Glu Arg Ile
        355                 360                 365

ccg aac ctc gaa ctc gcc gtg ccc acc gac ggg ttg gac atc aag tac  1152
Pro Asn Leu Glu Leu Ala Val Pro Thr Asp Gly Leu Asp Ile Lys Tyr
    370                 375                 380

gac ggc gtg ctc tac ggc ctg aac gag ctg ccc gtc cgc tgg tag      1197
Asp Gly Val Leu Tyr Gly Leu Asn Glu Leu Pro Val Arg Trp
385                 390                 395


<210> 8
<211> 399
<212> PRT
<213> Streptomyces natalensis

<400> 8
Met Thr Tyr Thr Asp Pro Ala Ala Pro Glu Thr Asp Pro Pro Ala Val
  1               5                  10                  15

Asp Phe Pro Gln Arg Lys Pro Gly Val Pro Phe Pro Pro Pro Asp Tyr
```

```
                20                  25                  30

Ala Asp Tyr Arg Asp Arg Lys Gly Leu Val Leu Ser Gln Leu Ser Asp
            35                  40                  45

Gly Lys Arg Val Trp Leu Val Thr Arg His Glu Asp Val Arg Ala Val
        50                  55                  60

Leu Thr Ser Pro Ser Ile Ser Ser Asn Pro Glu His Lys Gly Phe Pro
65                  70                  75                  80

Asn Val Gly Asn Leu Gly Val Pro Lys Gln Asp Gln Ile Pro Gly Trp
                85                  90                  95

Phe Val Gly Met Asp Ser Pro Glu His Asp Arg Phe Arg Lys Ala Leu
            100                 105                 110

Ile Pro Glu Phe Thr Val Arg Arg Val Arg Ala Met Lys Pro Ala Ile
        115                 120                 125

Glu Arg Thr Val Asp Ala Gln Leu Asp Ala Met Leu Ala Ala Gly Asn
    130                 135                 140

Thr Ala Asp Leu Val Ala Asp Phe Ala Leu Pro Ile Pro Ser Leu Val
145                 150                 155                 160

Ile Ser Ala Leu Leu Gly Val Pro Pro Ala Asp Arg Glu Phe Phe Glu
                165                 170                 175

Ser Arg Thr Arg Val Leu Val Ser Leu Arg Ser Ser Thr Asp Asp Asp
            180                 185                 190

Arg Met Ala Ala Ala Lys Asp Leu Leu Arg Tyr Ile Asn Arg Leu Val
        195                 200                 205

Glu Ile Lys Gln Lys Trp Gly Gly Asp Asp Leu Ile Thr Arg Leu Leu
    210                 215                 220

Ala Thr Gly Ala Ile Ala Pro His Glu Met Ser Gly Val Leu Met Leu
225                 230                 235                 240

Leu Leu Ile Ala Gly His Glu Thr Thr Ala Asn Asn Ile Ala Leu Gly
                245                 250                 255

Val Val Thr Leu Leu Ala Asn Pro Gln Trp Ile Gly Asp Asp Arg Ala
            260                 265                 270

Val Glu Glu Thr Leu Arg Phe His Ser Val Ala Asp Leu Val Ser Leu
        275                 280                 285

Arg Val Ala Val Gln Asp Val Glu Ile Ala Gly Gln Leu Ile Lys Ala
    290                 295                 300

Gly Glu Gly Ile Val Pro Leu Val Ala Ala Ala Asn His Asp Glu Asn
305                 310                 315                 320

Ala Phe Glu Cys Pro His Ala Phe Asp Pro Ser Arg Ser Ala Arg His
                325                 330                 335

His Val Ala Phe Gly Tyr Gly Val His Gln Cys Leu Gly Gln Asn Leu
            340                 345                 350

Val Arg Ile Glu Met Glu Val Ala Tyr Arg Lys Leu Phe Glu Arg Ile
        355                 360                 365
```

```
Pro Asn Leu Glu Leu Ala Val Pro Thr Asp Gly Leu Asp Ile Lys Tyr
    370                 375                 380

Asp Gly Val Leu Tyr Gly Leu Asn Glu Leu Pro Val Arg Trp
385                 390                 395
```

<210> 9
<211> 1194
<212> DNA
<213> Streptomyces natalensis

<220>
<221> CDS
<222> (1)..(1194)
<223> ORF3

<400> 9

```
atg acc gcc gcc tcc cac gac ctg ccc tgc ctc aac ctc gaa ccg ccc   48
Met Thr Ala Ala Ser His Asp Leu Pro Cys Leu Asn Leu Glu Pro Pro
  1               5                  10                  15

aaa atg ctg aaa ctg agc ccg ctg ctg cgc gcc ttg cag gac cgg ggg   96
Lys Met Leu Lys Leu Ser Pro Leu Leu Arg Ala Leu Gln Asp Arg Gly
            20                  25                  30

ccg atc cac cgg gtg cgc aca ccc gcc ggg gac gag gcg tgg ctg gtg   144
Pro Ile His Arg Val Arg Thr Pro Ala Gly Asp Glu Ala Trp Leu Val
        35                  40                  45

acc cgc cac gcc gag ctc aag cag ctg ctg cac gac gag cgc atc ggc   192
Thr Arg His Ala Glu Leu Lys Gln Leu Leu His Asp Glu Arg Ile Gly
    50                  55                  60

cgc acg cac ccc gac ccg ccc tcc gcc gcc cag tac gta cgc agc ccc   240
Arg Thr His Pro Asp Pro Pro Ser Ala Ala Gln Tyr Val Arg Ser Pro
65                  70                  75                  80

ttc ctg gac ctg ctg atc agc gac gcc gac gcc gag tcc ggg cgt cgg   288
Phe Leu Asp Leu Leu Ile Ser Asp Ala Asp Ala Glu Ser Gly Arg Arg
                85                  90                  95

cag cac gcc gag acc cgc cgc ctg ctc act ccg ttg ttc tcg gcc cgg   336
Gln His Ala Glu Thr Arg Arg Leu Leu Thr Pro Leu Phe Ser Ala Arg
            100                 105                 110

cgc gtt ctg gaa atg cag ccg aag gtg gag gag gcc gcg gac acc ctg   384
Arg Val Leu Glu Met Gln Pro Lys Val Glu Glu Ala Ala Asp Thr Leu
        115                 120                 125

ctg gac gcg ttc atc gcc cag ggg cct ccc ggc gac ctg cac ggc gag   432
Leu Asp Ala Phe Ile Ala Gln Gly Pro Pro Gly Asp Leu His Gly Glu
    130                 135                 140

ctc acc gtg ccg ttc gcc ctc acg gtc ctc tgc gag gtc atc ggc gtg   480
Leu Thr Val Pro Phe Ala Leu Thr Val Leu Cys Glu Val Ile Gly Val
145                 150                 155                 160

ccg ccg cag cgc cgc gcg gag ctg acc aca ctg ctg gcc ggt atc gcc   528
Pro Pro Gln Arg Arg Ala Glu Leu Thr Thr Leu Leu Ala Gly Ile Ala
                165                 170                 175

aag ctg gac gac cgc gag ggc gcc gta cgg gca cag gac gac ctg ttc   576
Lys Leu Asp Asp Arg Glu Gly Ala Val Arg Ala Gln Asp Asp Leu Phe
            180                 185                 190
```

```
ggg tac gtg gca ggg ctg gtc gag cac aag cgg gcc gag ccc ggc cca   624
Gly Tyr Val Ala Gly Leu Val Glu His Lys Arg Ala Glu Pro Gly Pro
        195                 200                 205

gac atc atc tcc cgg ctg aac gac ggc gag ctg acc gag gac cgc gtg   672
Asp Ile Ile Ser Arg Leu Asn Asp Gly Glu Leu Thr Glu Asp Arg Val
    210                 215                 220

gca cac ctg gcc atg ggc ctg ctg ttc gcc ggg ctg gac agc gtc gcg   720
Ala His Leu Ala Met Gly Leu Leu Phe Ala Gly Leu Asp Ser Val Ala
225                 230                 235                 240

agc atc atg gac aac ggg gtg gtg ctg ctg gcc gcc cac ccc gat cag   768
Ser Ile Met Asp Asn Gly Val Val Leu Leu Ala Ala His Pro Asp Gln
                245                 250                 255

cgc gcg gcg gcg ctg gcc gac ccc gac gtg atg gcg cgt gcc gtg gag   816
Arg Ala Ala Ala Leu Ala Asp Pro Asp Val Met Ala Arg Ala Val Glu
            260                 265                 270

gag gtg ctg cgg acc gcc cgg gcc ggc ggg tcg gtc ctg ccg ccg cgc   864
Glu Val Leu Arg Thr Ala Arg Ala Gly Gly Ser Val Leu Pro Pro Arg
        275                 280                 285

tac gcc agc gag gac atg gaa ttc ggc ggg gtg acg ata cgg gcc gga   912
Tyr Ala Ser Glu Asp Met Glu Phe Gly Gly Val Thr Ile Arg Ala Gly
    290                 295                 300

gac ctg gtc ctg ttc gac ctc ggc ctg ccc aac ttc gac gag cgg gcg   960
Asp Leu Val Leu Phe Asp Leu Gly Leu Pro Asn Phe Asp Glu Arg Ala
305                 310                 315                 320

ttc aca ggg ccg gag gaa ttc gac gcc gcc agg acc ccc aat ccc cat   1008
Phe Thr Gly Pro Glu Glu Phe Asp Ala Ala Arg Thr Pro Asn Pro His
                325                 330                 335

ctg acc ttc ggc cac ggc atc tgg cac tgc atc ggc gcc ccc ctc gcg   1056
Leu Thr Phe Gly His Gly Ile Trp His Cys Ile Gly Ala Pro Leu Ala
            340                 345                 350

cgc ctg gaa ctc agg acg atg ttc acc aag ctg ttc acc cgc ctg ccg   1104
Arg Leu Glu Leu Arg Thr Met Phe Thr Lys Leu Phe Thr Arg Leu Pro
        355                 360                 365

gaa ctg cgc ccg gaa ctt ccg gtg gag caa ctg cgc ctg aag gag ggc   1152
Glu Leu Arg Pro Glu Leu Pro Val Glu Gln Leu Arg Leu Lys Glu Gly
    370                 375                 380

cag ctg tcg ggc ggc ttc gcc gag ctc cgg gtg gtc tgg tag           1194
Gln Leu Ser Gly Gly Phe Ala Glu Leu Arg Val Val Trp
385                 390                 395
```

<210> 10
<211> 398
<212> PRT
<213> Streptomyces natalensis

<400> 10

```
Met Thr Ala Ala Ser His Asp Leu Pro Cys Leu Asn Leu Glu Pro Pro
  1               5                  10                  15

Lys Met Leu Lys Leu Ser Pro Leu Leu Arg Ala Leu Gln Asp Arg Gly
            20                  25                  30
```

Pro Ile His Arg Val Arg Thr Pro Ala Gly Asp Glu Ala Trp Leu Val
35 40 45

Thr Arg His Ala Glu Leu Lys Gln Leu Leu His Asp Glu Arg Ile Gly
50 55 60

Arg Thr His Pro Asp Pro Pro Ser Ala Ala Gln Tyr Val Arg Ser Pro
65 70 75 80

Phe Leu Asp Leu Leu Ile Ser Asp Ala Asp Ala Glu Ser Gly Arg Arg
85 90 95

Gln His Ala Glu Thr Arg Arg Leu Leu Thr Pro Leu Phe Ser Ala Arg
100 105 110

Arg Val Leu Glu Met Gln Pro Lys Val Glu Glu Ala Ala Asp Thr Leu
115 120 125

Leu Asp Ala Phe Ile Ala Gln Gly Pro Pro Gly Asp Leu His Gly Glu
130 135 140

Leu Thr Val Pro Phe Ala Leu Thr Val Leu Cys Glu Val Ile Gly Val
145 150 155 160

Pro Pro Gln Arg Arg Ala Glu Leu Thr Thr Leu Leu Ala Gly Ile Ala
165 170 175

Lys Leu Asp Asp Arg Glu Gly Ala Val Arg Ala Gln Asp Asp Leu Phe
180 185 190

Gly Tyr Val Ala Gly Leu Val Glu His Lys Arg Ala Glu Pro Gly Pro
195 200 205

Asp Ile Ile Ser Arg Leu Asn Asp Gly Glu Leu Thr Glu Asp Arg Val
210 215 220

Ala His Leu Ala Met Gly Leu Leu Phe Ala Gly Leu Asp Ser Val Ala
225 230 235 240

Ser Ile Met Asp Asn Gly Val Val Leu Leu Ala Ala His Pro Asp Gln
245 250 255

Arg Ala Ala Ala Leu Ala Asp Pro Asp Val Met Ala Arg Ala Val Glu
260 265 270

Glu Val Leu Arg Thr Ala Arg Ala Gly Gly Ser Val Leu Pro Pro Arg
275 280 285

Tyr Ala Ser Glu Asp Met Glu Phe Gly Gly Val Thr Ile Arg Ala Gly
290 295 300

Asp Leu Val Leu Phe Asp Leu Gly Leu Pro Asn Phe Asp Glu Arg Ala
305 310 315 320

Phe Thr Gly Pro Glu Glu Phe Asp Ala Ala Arg Thr Pro Asn Pro His
325 330 335

Leu Thr Phe Gly His Gly Ile Trp His Cys Ile Gly Ala Pro Leu Ala
340 345 350

Arg Leu Glu Leu Arg Thr Met Phe Thr Lys Leu Phe Thr Arg Leu Pro
355 360 365

Glu Leu Arg Pro Glu Leu Pro Val Glu Gln Leu Arg Leu Lys Glu Gly

370 375 380

Gln Leu Ser Gly Gly Phe Ala Glu Leu Arg Val Val Trp
385 390 395

<210> 11
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Forward primer
ermE promoter

<400> 11
aaactgcagc tctagaggcg gcttgcgccc gatgctagtc 40

<210> 12
<211> 70
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Reverse primer
ermE promoter

<400> 12
aaactgcagc tctagatggc ccgggtatcg atcgtcgacg gcatgcggat cctaccaacc 60

ggcacgattg 70

Claims

1. A nucleotide sequence comprising SEQ ID NO. 5, SEQ ID NO. 7 or SEQ ID NO. 9, and homologues or fragments thereof.

2. A nucleotide sequence as depicted by SEQ ID NO. 5, SEQ ID NO. 7 or SEQ ID NO. 9, and homologues thereof.

3. An enzyme obtainable by expressing a nucleotide sequence according to claims 1-2 in a micro-organism chosen from the group of *Streptomyces* species and heterologous species.

4. An enzyme comprising the amino acid sequence as depicted by SEQ ID NO. 6, SEQ ID NO. 8, or SEQ ID NO. 10, and homologues and fragments thereof.

5. A method for overexpressing a gene encoding an enzyme according to claim 3 or 4 in *Streptomyces* by attaching a promoter sequence to the gene, transferring the promoter-gene complex into a cell and bringing the gene to expression.

6. A method for knocking out a gene encoding an enzyme according to claim 3 or 4 in *Streptomyces* by disrupting the coding sequence of the gene.

7. A method for expressing a gene encoding an enzyme according to claim 3 or 4, in a heterologous micro-organism by attaching a promoter sequence to the gene, transferring the promoter-gene complex into a cell and bringing the gene to expression.

8. A recombinant microorganism strain obtainable by a method according to claims 5-7.

9. A method for preparing pimaricin involving the use of a recombinant *Streptomyces* obtainable by a method according to claim 5.

10. A method according to claim 9, wherein *Streptomyces* is *Streptomyces natalensis*.

11. A method for preparing a biomolecule of interest involving the use of a recombinant *Streptomyces* obtainable by a method according to claim 6 or 7.

12. A biomolecule of interest obtainable by a method according to claim 11.

EP-2959 DSM N.V.

GENES ENCODING ENZYMES IN THE BIOSYNTHESIS OF PIMARICIN AND THE APPLICATION THEREOF

ABSTRACT

The invention relates to genes encoding enzymes which are fundamental in the biosynthesis of pimaricin, and to the enzymes which are encoded by said genes. The invention further relates the application of said genes for modifying the biosynthesis of pimaricin, as well as for the biosynthesis of new compounds.

Figure 1

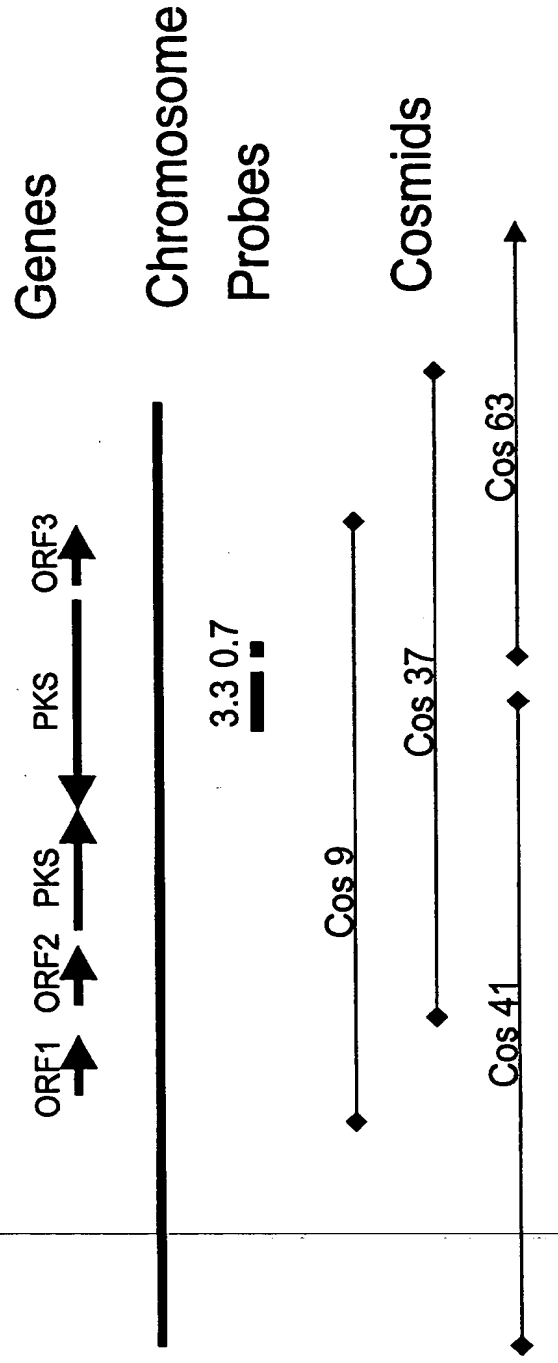
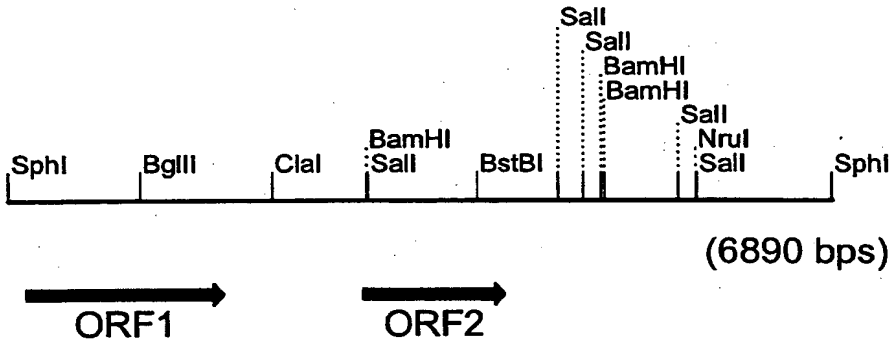

Genes
Chromosome
Probes
Cosmids
ORF3
PKS
PKS
ORF2
ORF1
3.3 0.7
Cos 9
Cos 37
Cos 63
Cos 41

Figure 2


SalI
SalI
BamHI
BamHI
SalI
NruI
SalI
BamHI
SalI
SphI
BglII
ClaI
BstBI
SphI
(6890 bps)
ORF1
ORF2

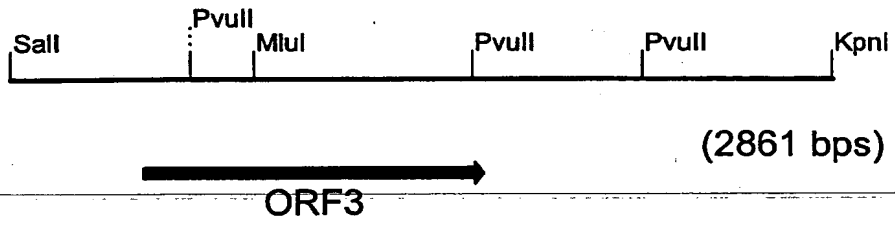
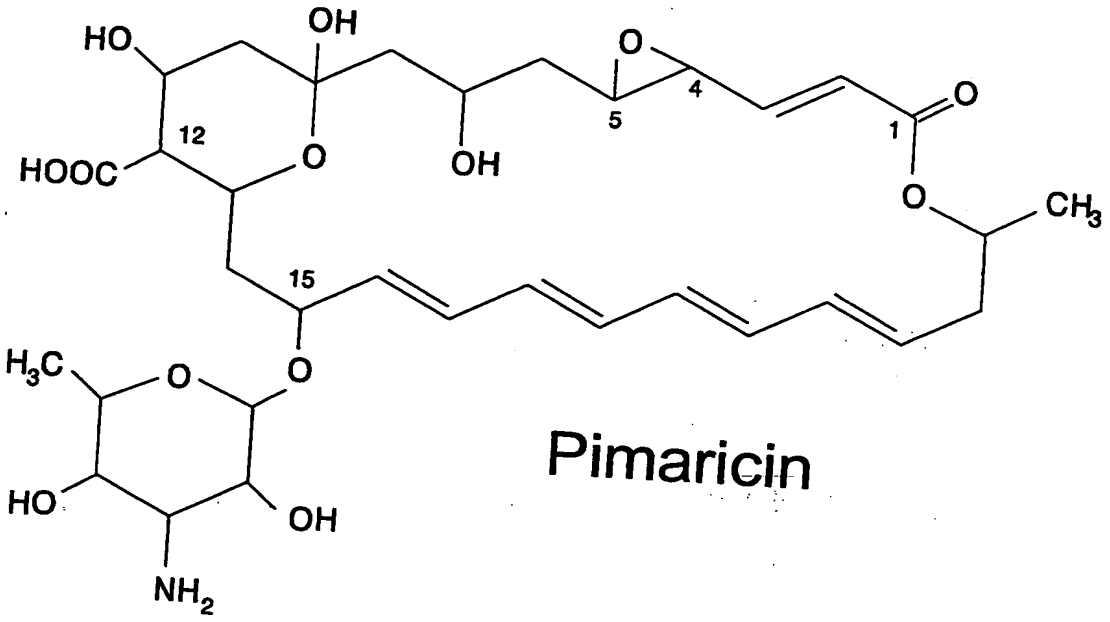
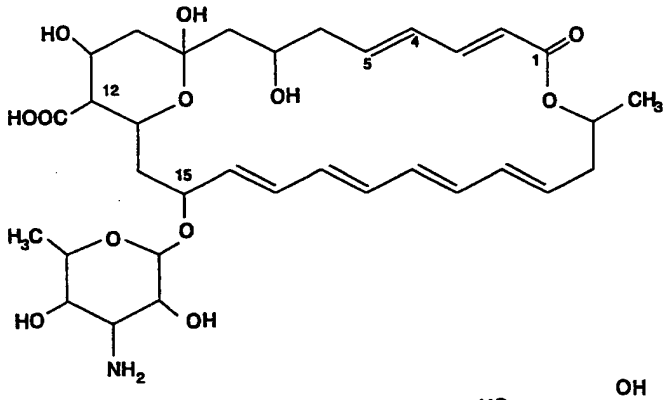

PvuII
SalI
MluI
PvuII
PvuII
KpnI
(2861 bps)
ORF3

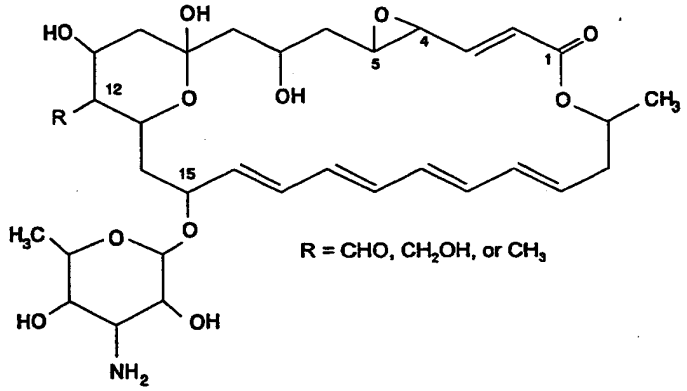
Figure 3a



Pimaricin

Figure 3b


HO
OH
HOOC
12
O
OH
5
4
1
O
O
CH3
15
O
H3C
O
HO
OH
NH2


HO
OH
O
4
5
1
O
12
R
O
OH
O
CH3
15
O
H3C
O
HO
OH
NH2
R = CHO, CH2OH, or CH3

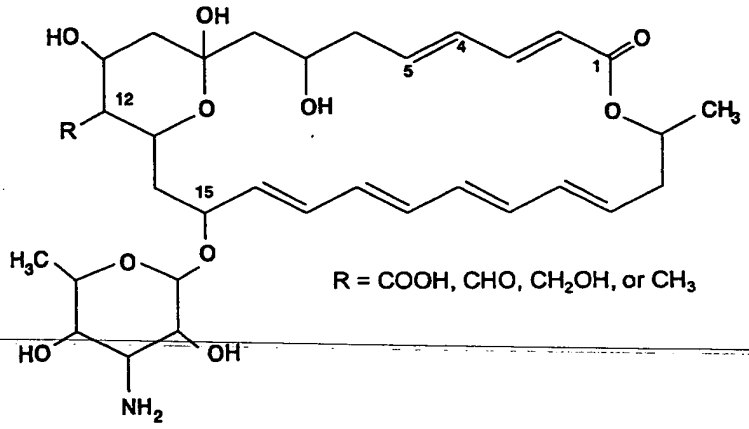
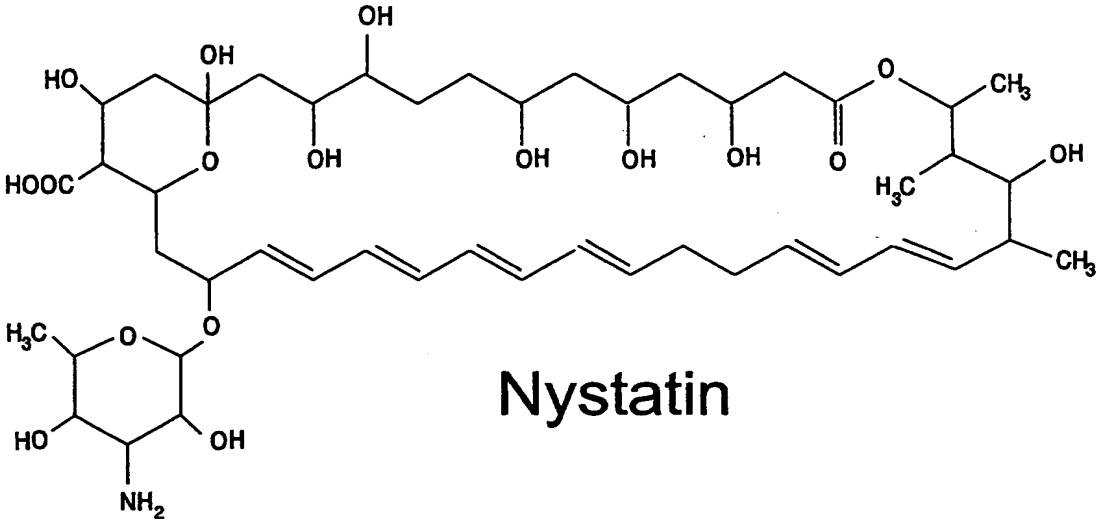
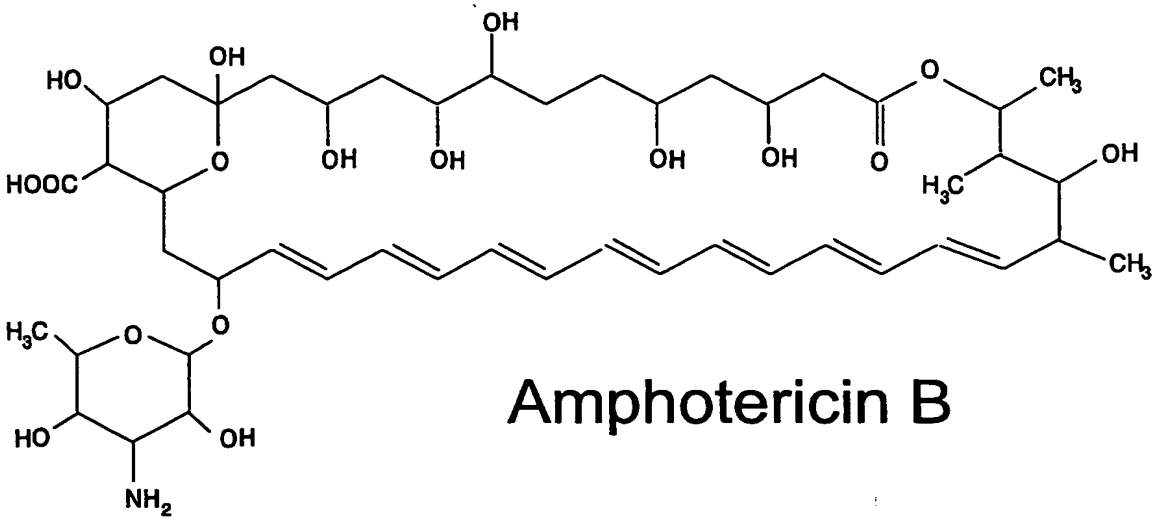

HO
OH
5
4
1
O
12
R
O
OH
O
CH3
15
O
H3C
O
HO
OH
NH2
R = COOH, CHO, CH2OH, or CH3

Figure 4



Nystatin



Amphotericin B